Journal of the
OPTICAL SOCIETY
of AMERICA

VOLUME 36, NUMBER 5 MAY, 1946

The Concentrated-Arc Lamp

W. D. BUCKINGHAM AND C. R. DEIBERT
The Western Union Telegraph Company, Electronics Division, Water Mill, New York
(Received February 27, 1946)

DURING the war a new type of light source was developed under a contract issued through the Optics Division of the National Defense Research Committee. The new lamp, called the concentrated-arc lamp, is basically a direct-current arc lamp made with permanent, fixed electrodes which are sealed into a glass bulb filled with an inert gas. The name, "concentrated-arc," comes from a characteristic of the lamp which makes it possible to concentrate the arc activity upon a small portion of the electrode so as to produce a very high intensity light source in the form of a luminous circular spot which is fixed in position, sharply defined and uniformly brilliant.

The degree to which this concentration may be carried is illustrated by the 2-watt concentrated-arc lamp whose luminous spot is only 0.003 of an inch in diameter with a maximum brightness of 100 candles per sq. mm. Ordinary tungsten filament lamps operate at about 10 candles per square millimeter while the brightness of the crater of a carbon arc lamp is about 135 candles per square millimeter. The area of the spot formed increases with the current so that a 100-watt lamp has a spot 0.059 of an inch in diameter, while the spot of a 1500-watt lamp is 0.35 inch in diameter.

The source of the light is an incandescent spot which forms on the specially prepared negative electrode or cathode. This cathode is the unique element of the new lamp. It is made by packing zirconium oxide into a small cup or the open end of a tube which is made of tungsten, molybdenum, or tantalum, these metals being selected because of their high melting temperatures.

The positive electrode or anode, also made of a metal with a high melting point, consists of a simple sheet or plate which has sufficient radiating surface so that during operation it will reach no more than a dull red heat.

These two electrodes are mounted in the bulb so that the exposed oxide surface of the cathode is but a few hundredths of an inch from and directly behind a hole in the center of the anode. This hole is slightly larger in diameter than the cathode tube or cup and provides a window for the emergence of light from the cathode.

After the bulb has been evacuated, it is filled with an inert gas, usually argon, to almost atmospheric pressure. The cathode is then put through a "forming" process. To do this a high potential direct-current source, with suitable current limiting resistors in series, is connected to the electrodes so that an arc strikes between the anode and the metallic tube of the cathode. After a few seconds, the cathode tube becomes red hot and heats the zirconium oxide packed in it to a temperature where the oxide becomes electrically conductive. The arc then strikes between the anode and the oxide and ionic bombardment raises the temperature of the surface

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

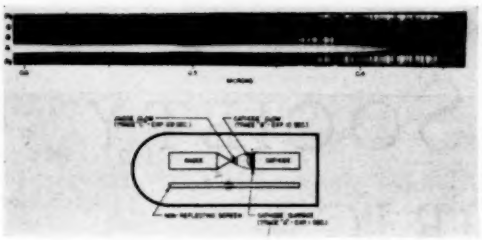


FIG. 1. Spectrogram of various regions of a concentrated arc.

of the oxide to or above its melting point of 3000° K. The molten oxide flows and bonds itself to the sides of the metal tube forming a smooth glassy surface.

The exact mechanism of the next phase of the forming sequence is not fully understood but it is thought that in the molten state and under the intense ionic bombardment of the arc, some of the zirconium oxide is reduced to metallic zirconium. The zirconium atoms form a thin surface layer in equilibrium with the molten oxide. The released oxygen migrates to the anode which always shows signs of oxidation.

Zirconium metal is apparently a better electron emitter at high temperatures than is the oxide and it also has a lower melting temperature; thus as soon as the zirconium surface layer is formed the temperature of the cathode drops slightly and the underlying oxide solidifies and supports the thin film of molten metal on its surface. It is this film of molten metal which is the chief source of the visible radiation from the lamps. The film, once formed during manufacture, remains to be heated and become incandescent whenever the lamp is relighted. It is so thin that surface tension holds it to the oxide backing so the lamps may be burned in any position.

Spectrograms of the portion of the arc stream very near the cathode always show zirconium lines. This indicates that some evaporation of free zirconium occurs. Very little zirconium is lost however since positive zirconium ions are produced by electron bombardment and are driven back to the cathode. If any metal atoms do escape from the cathode, they are replaced by reduction of the underlying oxide. As a result of these processes, the lamps have lives which are measured in hundreds of hours.

The cathode-current densities in the Concentrated-Arc Lamp vary from about 250 amperes per square centimeter for the 100-watt lamp to about 900 amperes per square centimeter for the 2-watt lamp. Assuming the electron emitter to be the thin zirconium layer at a temperature slightly below 3000° K, the melting point of the oxide, and using the constants commonly given for a zirconium filament in a vacuum, values of electron emission are obtained which are of the order of 500 amperes per square centimeter. This value is entirely in line with the actual current densities found in the lamps and seems to confirm the present belief that the arc is maintained entirely by thermionic emission, and that the active surface of the cathode consists of a thin layer of zirconium atoms at a temperature considerably above the normal melting point of the bulk metal.

The radiation from the concentrated-arc lamp appears to be divided into three parts as follows:

1. Continuous radiation from the molten cathode surface.
2. Line radiation from the excited gas and vapor.
3. Continuous radiation in the spectral region from at least 3500A to 5000A originating in the excited gas and vapor.

The existence of the three types of radiation is shown by the spectrograms in Fig. 1. The three exposures were made from the cathode spot, cathode glow, and anode glow portions of the arc of a specially constructed 100-watt lamp. The cathode spot is by far the brightest portion. Even though the exposure times for the three

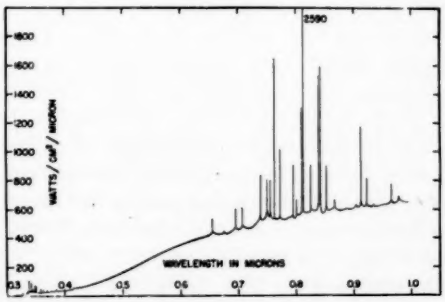


FIG. 2. Spectral distribution of radiation from a 100-watt concentrated-arc lamp.

traces varied in the ratio of 1:10:100, it is found that the trace of the cathode spot area is more than ten times as intense as that of the cathode glow area.

Thus the molten incandescent cathode surface or cathode spot furnishes the greatest amount of radiation from the arc. It is, in all probability, thermal radiation from the bound molecules and atoms present in the surface. The spectral distribution of this radiation is essentially that of a gray body.

The characteristic radiation of the rare gas filling and the zirconium vapor comes from the arc stream and in particular from the cathode glow region.

Spectrograms of 100-watt concentrated-arcs in argon show lines of excited and singly ionized argon and zirconium atoms and doubly ionized zirconium atoms (AI, ZrI, AII, ZrII, and ZrIII). The excited argon, AI, and the singly ionized zirconium, ZrII, atomic lines are the strongest and most numerous, only a few very weak AII and ZrIII lines appearing.

The continuum shown in the spectrogram of the cathode glow region seems to have its origin in this dense cloud of argon gas and zirconium ions which is very close to the molten cathode surface. Since special precautions were taken in constructing the lamp to make sure that scattered or reflected light from the molten cathode surface could not enter the spectrograph when the cathode glow was used as the source of radiation, it seems quite certain that this continuum has a different origin than that of the cathode spot.

Thus concentrated-arc lamps emit radiation from two main sources, the white hot zirconium cathode surface and the cloud of excited zirconium vapor and argon gas extending for a few thousandths of an inch from the cathode. The portion which originates from the cathode surface has a continuous spectral distribution. It extends in measurable amounts from 2500A in the ultraviolet, through the visible, reaching a maximum near 10,000A and on into the infrared. That portion of the radiation which comes from the cloud of excited vapor and gas shows three principal spectra, a continuum extending from the ultraviolet to about 5000A, the normal, singly and doubly ionized zirconium spectrum and the normal and singly ionized argon spectrum.

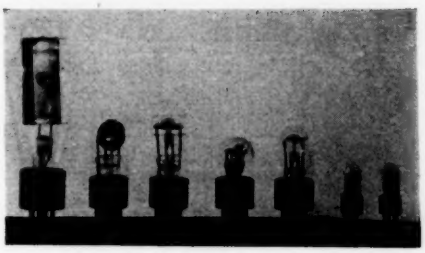


FIG. 3. 100-, 25-, 10-, and 2-watt concentrated-arc lamps.

The spectral distribution characteristic shown in Fig. 2 thus represents the combination or sum of these several individual spectra. Radiation shorter than 3000A or longer than 5 microns is not transmitted by the type of glass used for the bulbs of standard type lamps.

A line of standard size lamps has been developed in sizes ranging from 2 to 100 watts. Pictures of these lamps are shown in Fig. 3. Larger lamps are made but designs have not been standardized.

The diameter of the cathode spot of a given lamp depends upon the current. If the current is increased, the spot slowly grows larger, taking several seconds to adjust itself to the new condition. Figure 4 shows how the diameter of the light sources vary from 0.05 mm to 3.5 mm as the lamp currents are changed in the several standard sizes of lamps. While the lamps are designed to operate at a definite current value, it is possible to adjust the spot size by changing the current. The upper limit of this process is reached when the spot completely covers the activated zirconium oxide surface. If the current is increased further, the spot becomes brighter but not larger and the life of the lamp is shortened. When the current is very much less than normal, the spot becomes variable in size and position and the lamp is unstable.

One of the advantages of the concentrated arc is its high brightness. As shown by Fig. 5, the maximum brightness of standard lamps at their normal operating current varies between 40 and 100 candles per square millimeter.

To the eye, the cathode spot appears to have a uniform and constant brilliance. Measure-

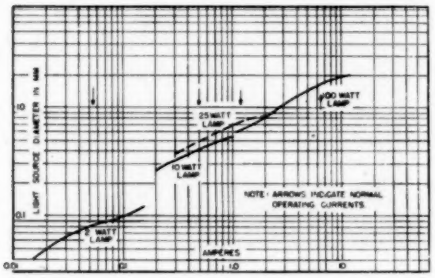


FIG. 4. Change of light source diameter with current of concentrated-arc lamp.

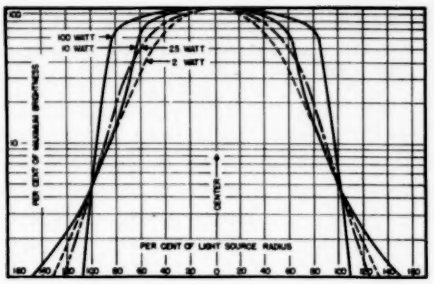


FIG. 6. Average cathode brightness distribution of concentrated-arc lamps.

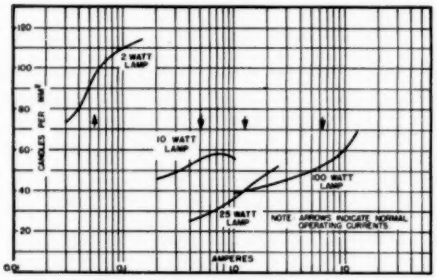


FIG. 5. Change of maximum brightness with current of concentrated-arc lamp.

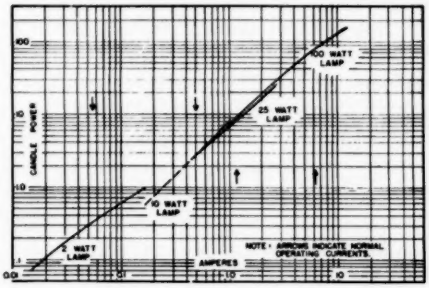


FIG. 7. Change of candlepower with current of concentrated-arc lamps.

ments show that the brightest part is near the center. The average brightness variation across the spot, for the several sizes of lamps, is shown in Fig. 6. Variations in the position taken by the arc stream and irregularities in the cathode surface may produce an unsymmetrical and changing brightness distribution. Also, the spot position may shift slowly during operation by an amount equal to a small percentage of its own diameter. These factors combine to produce variations of about ten percent in brightness and total emitted light.

The candlepower increases with the current, maintaining an almost linear relationship over a very wide range as shown by the curves of Fig. 7.

The efficiency of concentrated-arc lamps, as measured in candlepower per watt input to the lamp, varies between 0.15 for the 2-watt lamp to 0.8 for the 100-watt lamp. This characteristic is shown in Fig. 8. Comparable figures for tungsten filament lamps range from 0.54 for a 6-watt lamp to 1.29 for 100-watt lamps.

The spatial light distribution curve of Fig. 9 plots the light about the axis of the cathode and shows that concentrated-arc lamps have a cosine distribution.

The average changes in the major characteristics of a group of twenty-five 10-watt lamps during aging is shown by the curves of Fig. 10. These show that during the first few hours of running the candlepower and light spot diameter will decrease while the maximum brilliance increases. After about 100 hours of operation, these characteristics become reasonably stable around 80 percent of the initial candlepower, 75 percent of the initial light spot diameter, and 140 percent of the initial brilliance.

The mortality curve of Fig. 11 shows the average life of 2-watt lamps to be 175 hours. Similar data on larger lamps give 700, 800, and 1000 hours on 10-, 25-, and 100-watt lamps although individual lamps have shown lives up to 5000 hours. Failure is usually caused by loss or shrinkage of the cathode-filling material.

Concentrated-arc lamps have a negative volt-ampere characteristic as is shown by the curves of Fig. 12. Consideration must be given this fact in the design of their power supplies.

The lamps are started with a high voltage which breaks down the gap in the lamps. They are usually run from a rectifier or direct-current generator or battery with sufficient ballast resistance in series to limit the current to its normal value. Some lamps have been made with multiple anodes which permit their operation from an alternating current supply, in which case the lamps act as their own rectifiers.

Two-watt lamps require 1000 volts for starting and 130 volts or more for running. Larger lamps need up to 2000 volts for starting and may be run from sources supplying from 24 volts up, higher supply voltages resulting in more stable operation. The a.c. type of lamps are frequently started without the use of high voltages by the aid of an auxiliary tungsten filament which is built into these lamps to furnish the initial ionization required to break down the gap and establish the arc.

Power supplies, such as shown in Fig. 13, can be designed to operate from almost any power source and deliver the necessary starting and running voltages for any size lamp.

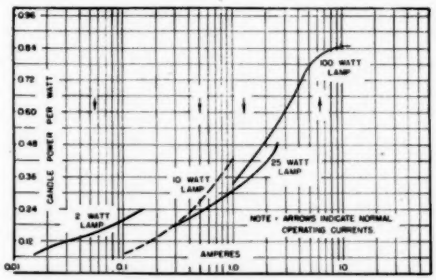


FIG. 8. Change of efficiency with current of concentrated-arc lamps.

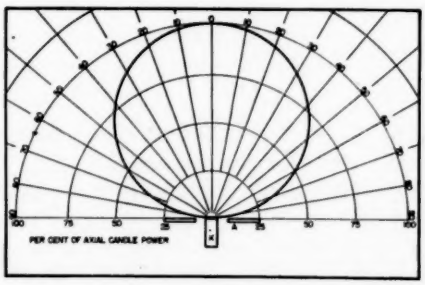


FIG. 9. Spatial distribution of radiation of concentrated-arc lamps.

The concentrated arc is therefore a new type of lamp which offers a unique combination of properties. The source of the radiation is a circular spot which is fixed in position, sharply defined, and uniformly brilliant. It is very bright, being several times brighter than tungsten filament lamps and almost as bright as the carbon arc. The source may be as small as 0.003 inch

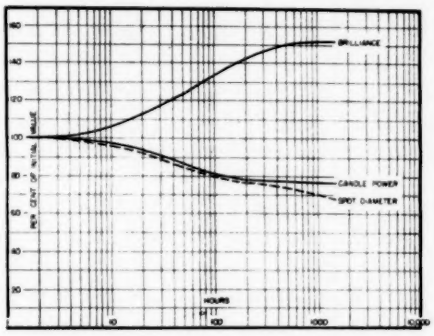


FIG. 10. Change of characteristics with age of 10-watt concentrated-arc lamps.

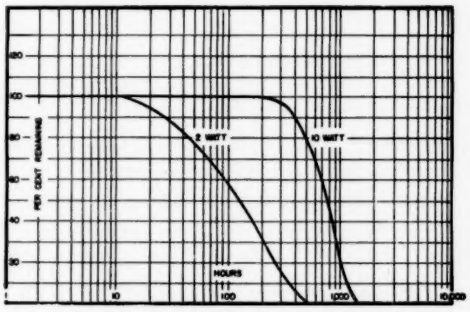


FIG. 11. Mortality curves for 2- and 4-watt concentrated-arc lamps.

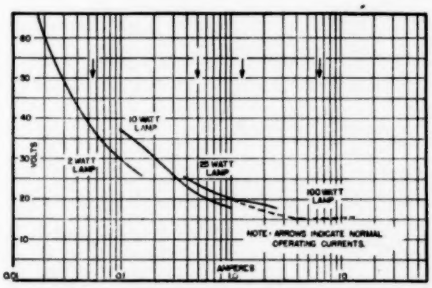


FIG. 12. Volt-ampere characteristics of concentrated-arc lamps.

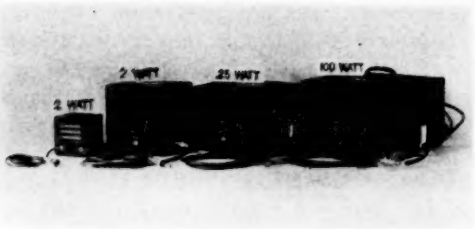


FIG. 13. A.c. operated power supplied for concentrated-arc lamps.

or as large as several tenths of an inch. The principle radiation is emitted as a continuum which extends well into the ultraviolet and infrared. The lamps start almost instantly and burn quietly without need of attention or adjustment. They may be burned in any position, are reasonably rugged and have lives which are measured in hundreds of hours. These properties make the concentrated-arc lamp useful for narrow beam and high intensity projection and illumination applications. It may also be used to furnish a high intensity continuum for absorption studies. In small sizes, the lamps are a close approach to a point source and have applications in optical testing and demonstrating, lensless projection and enlargement, photography, and in many entirely new devices which a brilliant point source will make possible.

ACKNOWLEDGMENT

The writers wish to express their appreciation to the members of Section 16.4 of the N.D.R.C., particularly Drs. O. S. Duffendack and Saul Dushman, also to Dr. W. S. Huxford of Northwestern University and Dr. H. L. Smith of Michigan State Normal College for their suggestions, help, and assistance in the development of this lamp.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 36, NUMBER 5 MAY, 1946

A Classification of Photographic Lens Types

R. KINGSLAKE
Eastman Kodak Company, Rochester, New York
(Received March 15, 1946)

A DETAILED classification of the various types of construction used in photographic objectives has long been needed, but previous attempts to develop such a classification using only a few clearly defined types have been unsatisfactory because of the large number of exceptions and borderline cases.

The classification described in this paper, except for some trifling modifications, has been in use in the Eastman Kodak Company for several years, and it has proved to be so satisfactory that its publication appears to be warranted in the hope that other lens designers and users may feel inclined to adopt it also. It has provided an excellent basis for the numbering of lens formulas; and by filing lens patents in accordance with the same system, any designer working on a given type of lens can at once draw out all the available patents relating to his particular problem.

The classification is based on the number of components in the lens, assuming that most photographic objectives fall into one or other of the following six broad groups:—

Singlet (single component with exterior stop) (Types *A*, *B*)

Duplet (two similar components about a central stop) (Types *D*, *E*, *F*)

Triplet (three airspaced components) (Types *G*, *H*, *J*, *K*)

Quadruplet (four airspaced components) (Types *L*, *M*, *N*)

Petzval (two dissimilar thin positive components widely separated) (Type *P*)

Telephoto (widely separated positive and negative components) (Type *T*)

It is surprising to find that perhaps 90 percent of all photographic objective designs do fall into the above simple classes. Furthermore, many of the remaining 10 percent of doubtful or borderline cases can usually be classified after a study of the published description of the design in the patent specification or elsewhere.

It is, however, useful to establish a "miscellaneous" classification for occasional odd types that bear no relation to any of the well-established classes. Also a few other special types such as reflecting systems, zoom lenses, anamorphosers, condensers, and viewfinders have been given distinguishing type letters that are outside the main classification given above.

It is realized, of course, that this system like any other is open to question. However, it has proved to be quite satisfactory and decidedly useful in actual practice, and it is presented here as a possible workable solution to the problem.

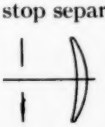
In the list that follows, each lens type is described in words and illustrated by an example taken from the literature. It must be remarked that the examples cited are not necessarily the first recorded uses of the various lens types, nor are they the most important instances, but they are typical representatives of their class. In some cases the trade names of examples are also included for a clearer understanding of the lens type, but the relationship of the patent cited to the actual lens named has not been checked. Most of the broader groups have been subdivided into sub-classes, illustrated by typical examples. Many small variations, such as the reversal of the order of crown and flint elements in a cemented doublet may occur without changing the lens type.

I wish to thank all my colleagues for their help and encouragement in the development of this classification and its reduction to a workable form. In particular, I am most indebted to Mr. F. E. Altman, Mr. F. M. E. Holmes, and Mr. H. F. Bennett of this organization for their kind criticism and very helpful suggestions.

PLAN OF THE CLASSIFICATION

A. Single Lenses

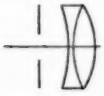
Front or rear single landscape lenses with a stop separated from the lens.



Wollaston, Phil. Trans. **2**, 370 (1812).

***B.** Singlet Lenses, compound*

Front or rear cemented (or closely airspaced) achromatic lens with an exterior stop. These systems are not intended to be used in pairs with a central stop, and are normally not spherically corrected.



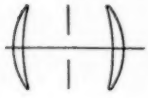
Chevalier, see von Rohr *Theorie und Geschichte des Photographischen Objectivs*, p. 94.

(The *B* type may be used to include telescope objectives if it is desired to do so.)

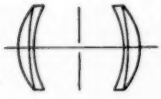
***C.** Condensers*

***D.** Duplets, symmetrical, not spherically corrected*

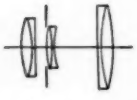
This class contains systems comprising two identical or closely similar singlets, symmetrically arranged about a central stop. It also includes the very old types consisting of three thin separated components, not spherically corrected.



D—(a) Simple Periscopic lenses. Steinheil, Brit. 2937/65 (Periskop).



D—(b) Achromatized systems. Harrison and Schnitzer, U. S. 35,605 (Globe Lens).



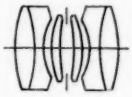
D—(c) Triplets, not spherically corrected. Dallmeyer (Triple Achromatic). See von Rohr *Theorie und Geschichte*, p. 170.

***E.** Duplets, symmetrical, spherically corrected and achromatic*

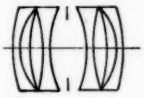
This group comprises symmetrical or almost symmetrical "double objectives," in which each half consists of a spherically corrected single cemented or closely airspaced unit.

E—(a) Two similar cemented doublets. Steinheil, U. S. 180,957 (Aplanat).

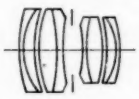
E—(b) Two similar cemented triplets. Von Höegh, U. S. 528,155 (Dagor).



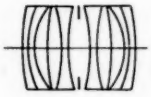
E—(c) Two similar triplets including narrow airspaces. Rudolph, Ger. 310,615 (Plasmat).



E—(d) Two similar cemented quadruplets. Rudolph, U. S. 532,398 (Double Protar, Series VIIa).



E—(e) Two similar quadruplets with airspaces. Lee, Brit. 376,044 (Taylor, Taylor & Hobson Convertible, Series XV).



E—(f) Two similar cemented quintuplets. Turner & Reich, U. S. 539,370.

E—(g) Others.

***F.** Duplets, unsymmetrical*

This class contains spherically and chromatically corrected duplets of which the two halves are decidedly dissimilar to one another.

Rudolph, U. S. 444,714 (Protar).

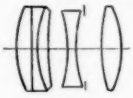
***G.** Triplets, simple*

These lenses contain three single airspaced elements, positive outside and negative inside.

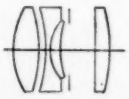
Taylor, U. S. 568,052 (Cooke lens).

***H.** Triplets, compound, i.e., containing one or more compound components*

Here *l* refers to a single element and *c* to a compound component. The compound components generally comprise one or more positive and negative elements cemented together or closely airspaced.

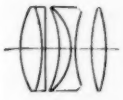


H—(a) Components c-l-l. Richter, U. S. 1,588,612.



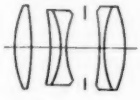
H—(b) Components l-c-l. Taylor, Brit. 22607/93.

H—(c) Components l-l-c. Rudolph, U. S. 721,240 (Tessar).

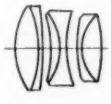


H—(d) Components c-c-l. Leitz, Ger. 526,308.

H—(e) Components c-l-c. Booth, U. S. 1,421,156 (Pentac).



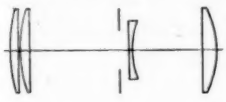
H—(f) Components l-c-c. Altman, U. S. 1,880,394.



H—(g) Components c-c-c. Berek, U. S. 1,939,098.

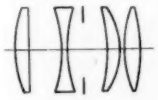
J. Triplets, divided, i.e., containing one or more divided components

In this group we find triplets of the *G* type in which one or more of the elements has been divided into two airspaced elements of the same sign as the original element.



J—(a) Front element divided. Sonnefeld, U. S. 1,825,828. (See also Type K, below)

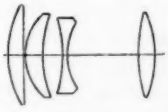
J—(b) Middle element divided. (These are included under Type L, below.)



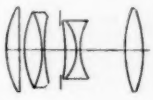
J—(c) Rear element divided. Bielicke, U. S. 1,540,752 (Tachar).

K. Triplets, with an additional meniscus component in the front airspace

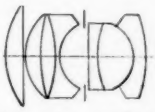
This class includes so many important modern objectives that we are justified in regarding it as a separate type, not merely a modified triplet as in class *J—(a)* above.



K—(a) The lenses in this group comprise four single elements, the second being meniscus and the third negative. Bertele, Brit. 237,212.



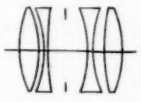
K—(b) Similar to K—(a) above, but having any or all of the four components compounded or split. Bertele, U. S. 1,584,272.



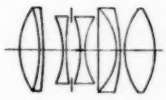
K—(c) Systems of this same type in which the second and third components are cemented together, with or without another glass element between them. Thus, in this type, the objective contains two positive exterior components with a thick meniscus-shaped negative component placed between them. Bertele, U. S. 1,975,678 (Sonnar).

L. Quadruplets

This class contains objectives having four airspaced components, in the order *plus—minus :minus—plus*, the inner negative components being either biconcave or meniscus with convex faces towards the stop.



L—(a) Four simple elements, with biconcave flints. Zschokke, U. S. 912,352 (Celor *f*/3.5).

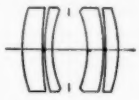


L—(b) As in L—(a), with some components compound or divided. Warmisham, U. S. 1,926,569.

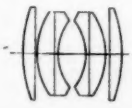
L—(c) One or both of the negative components is a meniscus, with the convex surface facing the stop. Rudolph, U. S. 1,565,205 (Kino Plasmat).

M. Quadruplets, meniscus

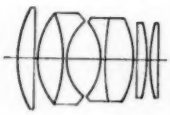
These are similar to the *L*-type, but the inner negative components are meniscus-shaped and concave to the stop.



M—(a) Four single elements. Meyer, Brit. 20349/00 (Aristostigmat).



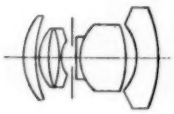
M—(b) One or both negative components cemented. Lee, Brit. 157,040 (Opic).



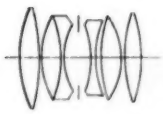
M—(c) As in M—(b), with the rear positive component divided. Lee, Brit. 373,950.

***N.** Quadruplets, unbalanced*

Here the components are not in the normal order; or the lens does not fall into either the *L* or *M* type.



N—(a) Components in the order plus—minus: plus—minus. Bertele, U. S. 2,084,309 (Biogon).

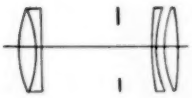


N—(b) Hybrid objectives in which one half is of the L-type and the other half of the M-type. Bielicke, U. S. 1,839,011 (Tachon).

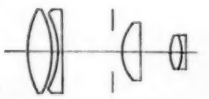
***P.** Petzval and Projection Types*

Two thin positive components widely separated, normally working at high aperture over a narrow field.

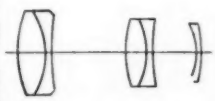
P—(a) Two separated cemented doublets. Repp, U. S. 1,479,251.



P—(b) Like P—(a), with one or both doublets air-spaced, or one or both doublets replaced by single elements. Petzval Portrait lens. See Richter, Merté, and von Rohr *Das Photographische Objectiv*, p. 324.

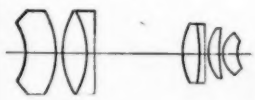


P—(c) More complicated types. Merté, U. S. 1,967,836 (*R*-Biotar).

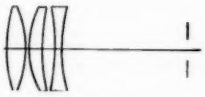


P—(d) With a field flattener. Wood, U. S. 2,076,190 (Kodascope Projector).

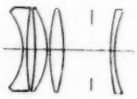
***Q.** Miscellaneous*



Q—(a) Microscope objectives. Boegehold, U. S. 2,206,155.



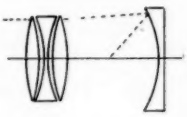
Q—(b) Profile projectors. Hopkins, Brit. 560,611.



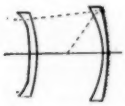
Q—(c) Unclassified photographic objectives. Taylor, Brit. 3799/12.

***R.** Mirror Lenses*

This class contains systems involving reflecting surfaces in addition to refracting surfaces.



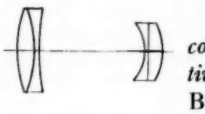
R—(a) Mirror-lens photographic objectives, using only spherical surfaces. Houghton, U. S. 2,350,112.



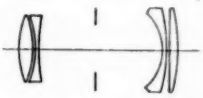
R—(b) Mirror lens systems involving aspheric surfaces (including Schmidt systems). Straubel, U. S. 2,170,979.

***T.** Telephoto Systems*

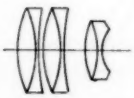
A telephoto lens is one in which the distance from the front vertex to the focal plane is less than the focal length. This result is achieved by means of a positive front member widely separated from a negative rear member.



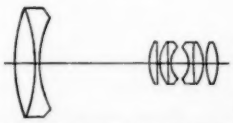
T—(a) Two thin cemented components, positive and negative, widely spaced. Booth, Brit. 139,719 (Dallon).



T—(b) Like T—(a), but one or both components air-spaced. Richter, U. S. 2,239,538 (Telikon).



T—(c) More elaborate systems. Cox, Brit. 552,024.

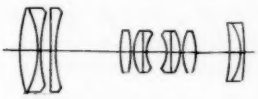


T—(d) Reversed Telephoto systems. In these the negative member is in front, and the back focus is equal to or greater than the focal length. Lee, U. S. 1,955,590.

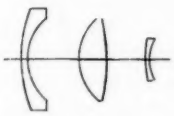
T—(e) Sky lenses. In these a field of 180° is covered, by the use of a highly distorting reversed telephoto construction. Hill, Brit. 225,398.

U. "Zoom" Lenses

These are lens systems of continuously variable focal length, in which the image is held constantly in focus by mechanical means.



U—(a) Complete zoom lens systems. Warmisham & Cisski, U. S. 1,947,669 (Varo).

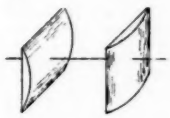


U—(b) Afocal zoom attachments to be added in front of a camera lens. Gramatzki, Brit. 449,434 (Transfokator).

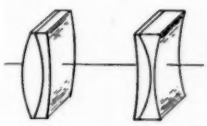
V. Viewfinders and Camera Rangefinders

W. Anamorphic Systems

These are photographic lens systems intended to give a greater magnification in one meridian than in the other, which effect may be produced either by cylindrical lens surfaces or by refracting prisms. This class does not include sound optics or scanning units (see *Z*).



W—(a) Complete anamorphic lens systems, using cylindrical lenses. Rudolph, Brit. 8512/98.



W—(b) Afocal anamorphic attachments, using cylindrical lenses. Newcomer, U. S. 1,945,951.

W—(c) Anamorphic attachments, using prisms. Phillips, Brit. 16923/06.

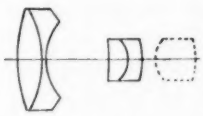
X. Lens Attachments

(excluding zoom or anamorphic attachments)

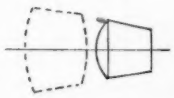
X—(a) Positive or "close-up attachments.

X—(b) Negative or "distance" attachments. Mengel, U. S. 1,169,731 (Distar).

X—(c) Soft focus, diffusing, and chromatic attachments. Weidert, U. S. 1,556,982 (Mollar).



X—(d) Afocal atachments to change the equivalent focal length of a lens. Bennett, U. S. 2,324,057.



X—(e) High aperture rear attachments. Luboshez, U. S. 1,952,268.

Y. Eyepieces and Loupes

Z. Miscellaneous Systems

Z—(a) Systems for recording, reproducing, and printing sound tracks on film; and scanning systems of all kinds. Maurer, U. S. 1,947,567.

Z—(b) Other systems.

Theory of an Efficient Interference Filter

ROBERT L. MOONEY
Department of Physics, Georgetown University, Washington, D. C.
(Received March 18, 1946)

1. INTRODUCTION

IN recent years it has become experimentally practicable to prepare high reflecting optical coatings which yield reflectivities in the visible region of the order of 50 percent. It is well known that the resolving power of the Fabry-Perot interferometer increases with the reflectivity of the metallic reflecting layer which is deposited on the surfaces of the plates. It is obvious that the efficiency of this type of interferometer, considered as a transmission instrument, decreases as the thicknesses of the reflecting metallic layers are increased to increase the reflectivities. The thicker layers of metal absorb so much energy that the usefulness of the instrument is seriously impaired, and the possibility of designing an efficient filter using silvered surfaces is precluded.

Exact formulas for the reflectivity and transmissivity of single and double layer coatings have been previously presented.[1] In the present paper a formula is derived for the transmissivity of an optically treated Fabry-Perot interferometer for the case of perpendicularly incident radiation and an approximate formula for the case of small angles of incidence. The predicted behavior of such an interferometer suggests the possibility of constructing an efficient filter by using two or more such interferometer units in series adjusted for different orders of interference. An example of such a filter is considered theoretically and the predicted results presented graphically. The primary purpose of this paper is to illustrate the fact that the method which treats the phenomenon of selective reflectivity and transmissivity of surface layers as a boundary problem in electromagnetic theory is straightforward and practically applicable to this type of problem which heretofore almost exclusively has been regarded as one to be solved by the superposition of amplitudes due to successive reflections.[2-6] This method becomes hopelessly complicated for cases involving as many as five layers—the number considered in this paper. In the present work, a striking symmetry appears in the mathematical relations and can be used as a guide throughout the derivation.

Previously published formulas[1] indicate that, with presently available materials, transmissivities as low as 65 percent—and lower—can be obtained for a definite wave-length using only two layers. Two successive transmissions through such layers with a non-phase-changing layer interposed will result in an over-all transmissivity of about 42 percent.

2. THEORETICAL DISCUSSION

Consider two optically plane slabs of glass coated with two transparent, homogeneous, and

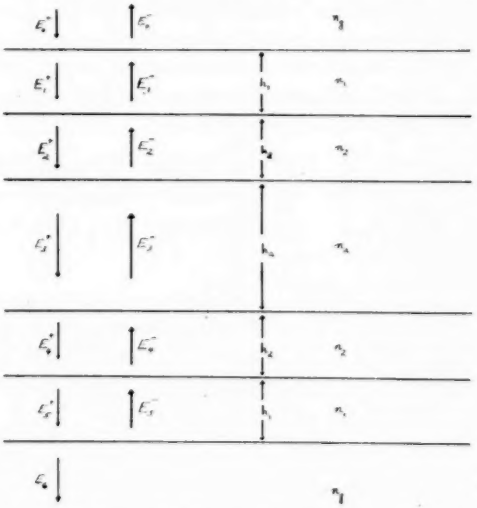


FIG. 1. Schematic diagram of interferometer showing directions of propagation of radiation of different amplitudes. Indices of refraction of the glass, air, and coating materials are indicated by n_g, n_a, n_1, and n_2.

[1] R. L. Mooney, J. Opt. Soc. Am. **35**, 574 (1945).

[2] K. B. Blodgett and I. Langmuir, Phys. Rev. **51**, 964 (1938).

[3] K. B. Blodgett, Phys. Rev. **57**, 921 (1940).

[4] A. F. Turner, J. App. Phys. **12**, 351 (1941).

[5] H. Bouasse and Z. Carriere, *Interferences* (Librarie Delagrave, Paris, 1923), p. 187.

[6] G. B. Airy, Trans. Camb. Phil. Soc. **4**, 479 (1830).

isotropic layers of uniform thickness. Let the index of refraction of the first layer—the one adjacent to the glass—be denoted by n_1, and let the index of the second layer be denoted by n_2. The indices of the glass and air will be denoted by n_g and n_a, respectively. Let the slabs be set up with their coated faces toward each other. The thickness of the layers of indices n_1 and n_2 will be denoted by h_1 and h_2, while the thickness of the air space will be denoted by h_a. Consider a beam of radiation of wave-length λ_a (in air) perpendicularly incident on the first layer from the glass of one of the slabs. The amplitude of the electric vector of this radiation is denoted by E_0^+. The amplitude of the electric vector of the radiation that is reflected from the system is denoted by E_0^-. The electric components of the radiation in the different layers are similarly denoted as illustrated in Fig. 1. For example, E_4^+ denotes the amplitude of the electric vector of the radiation that is propagated in the second (or top) layer of the second coated slab in the same direction as that of the initially incident radiation. The directions of the arrows in Fig. 1 indicate the directions of propagation of the different beams of radiation. The fundamental ideas necessary for setting up the equations which satisfy the boundary conditions of continuity of the tangential components of the electric and magnetic vectors of the radiation have been previously outlined.[1] By equating tangential components of the appropriate vectors on opposite sides of each bounding surface, one arrives at the following set of twelve simultaneous equations.

$$
\begin{aligned}
&E_0^+ + E_0^- = E_1^+ + E_1^-,\\
&n_g E_0^+ - n_g E_0^- = n_1 E_1^+ - n_1 E_1^-,\\
&E_1^+ \exp(ik_1h_1) + E_1^- \exp(-ik_1h_1) = E_2^+ \exp(ik_2h_1) + E_2^- \exp(-ik_2h_1),\\
&n_1 E_1^+ \exp(ik_1h_1) - n_1 E_1^- \exp(-ik_1h_1) = n_2 E_2^+ \exp(ik_2h_1) - n_2 E_2^- \exp(-ik_2h_1),\\
&E_2^+ \exp[ik_2(h_1+h_2)] + E_2^- \exp[-ik_2(h_1+h_2)]\\
&\qquad = E_3^+ \exp[ik_a(h_1+h_2)] + E_3^- \exp[-ik_a(h_1+h_2)],\\
&n_2 E_2^+ \exp[ik_2(h_1+h_2)] - n_2 E_2^- \exp[-ik_2(h_1+h_2)]\\
&\qquad = n_a E_3^+ \exp[ik_a(h_1+h_2)] - n_a E_3^- \exp[-ik_a(h_1+h_2)],\\
&E_3^+ \exp[ik_a(h_1+h_2+h_a)] + E_3^- \exp[-ik_a(h_1+h_2+h_a)]\\
&\qquad = E_4^+ \exp[ik_2(h_1+h_2+h_a)] + E_4^- \exp[-ik_2(h_1+h_2+h_a)],\\
&n_a E_3^+ \exp[ik_a(h_1+h_2+h_a)] - n_a E_3^- \exp[-ik_a(h_1+h_2+h_a)]\\
&\qquad = n_2 E_4^+ \exp[ik_2(h_1+h_2+h_a)] - n_2 E_4^- \exp[-ik_2(h_1+h_2+h_a)],\\
&E_4^+ \exp[ik_2(h_1+2h_2+h_a)] + E_4^- \exp[-ik_2(h_1+2h_2+h_a)]\\
&\qquad = E_5^+ \exp[ik_1(h_1+2h_2+h_a)] + E_5^- \exp[-ik_1(h_1+2h_2+h_a)],\\
&n_2 E_4^+ \exp[ik_2(h_1+2h_2+h_a)] - n_2 E_4^- \exp[-ik_2(h_1+2h_2+h_a)]\\
&\qquad = n_1 E_5^+ \exp[ik_1(h_1+2h_2+h_a)] - n_1 E_5^- \exp[-ik_1(h_1+2h_2+h_a)],\\
&E_5^+ \exp[ik_1(2h_1+2h_2+h_a)] + E_5^- \exp[-ik_1(2h_1+2h_2+h_a)] = E_6 \exp[ik_g(2h_1+2h_2+h_a)],\\
&n_1 E_5^+ \exp[ik_1(2h_1+2h_2+h_a)] - n_1 E_5^- \exp[-ik_1(2h_1+2h_2+h_a)]\\
&\qquad = n_g E_6 \exp[ik_g(2h_1+2h_2+h_a)].
\end{aligned}
\tag{1}
$$

Since both the incident and emergent radiation are in glass, an application of Poynting's Theorem shows that the transmissivity is given by

$$T = E_6 E_6^* / E_0 E_0^*. \tag{2}$$

If Eqs. (1) are solved for E_6, one obtains

$$E_6 = [64 n_1^2 n_2^2 n_a n_g / \Delta] E_0^+. \tag{3}$$

In evaluating the determinant, Δ, of the coefficients of the electric components in Eqs. (1)—regarding E_0^+ as known—the following abbreviations are useful:

$$
\begin{aligned}
&k_1h_1 = \phi_1,\\
&k_2h_2 = \phi_2,\\
&k_ah_a = \phi_a,\\
&B = (n_1-n_2)(n_1-n_g)\exp(i\phi_1) - (n_1+n_2)(n_1+n_g)\exp(-i\phi_1),\\
&C = (n_1+n_2)(n_1-n_g)\exp(i\phi_1) - (n_1-n_2)(n_1+n_g)\exp(-i\phi_1),\\
&K = C(n_2+n_a)\exp(i\phi_2) + B(n_2-n_a)\exp(-i\phi_2),\\
&D = C(n_2-n_a)\exp(i\phi_2) + B(n_2+n_a)\exp(-i\phi_2).
\end{aligned}
$$

Introducing these quantities into the factors

obtained in solving the equations, one obtains for the determinant

$$\Delta=[-K^2\exp(i\phi_a)+D^2\exp(-i\phi_a)]\times\exp i(2h_1+2h_2+h_a)k_g. \quad (4)$$

The transmissivity of the system is now given by

$$T=(64)^2n_2^4n_1^4n_a^2n_g^2/\Delta\Delta^*. \quad (5)$$

Equation (5) does not present a convenient form for the numerical calculation of the transmissivity. It will be observed that if the determinant be expressed as a multinomial in the exponentials, it may be written as the sum of eighteen terms. When this multinomial is multiplied by its complex conjugate, one obtains the square of the amplitude of the determinant as the sum of one-hundred-seventy-one terms. In this process the following abbreviations are convenient:

$$\begin{aligned}
N_1 &= -(n_1+n_2)^2(n_1-n_g)^2(n_2+n_a)^2,\\
N_2 &= -(n_1-n_2)^2(n_1+n_g)^2(n_2+n_a)^2,\\
N_3 &= 2(n_1^2-n_2^2)(n_1^2-n_g^2)(n_2+n_a)^2,\\
N_4 &= -(n_2-n_a)^2(n_1-n_2)^2(n_1-n_g)^2,\\
N_5 &= -(n_1+n_2)^2(n_1+n_g)^2(n_2-n_a)^2,\\
N_6 &= 2(n_1^2-n_2^2)(n_1^2-n_g^2)(n_2-n_a)^2,\\
N_7 &= -2(n_2^2-n_a^2)(n_1^2-n_2^2)(n_1-n_g)^2,\\
N_8 &= -2(n_2^2-n_a^2)(n_1^2-n_2^2)(n_1+n_g)^2,\\
N_9 &= 4(n_2^2-n_a^2)(n_1^2-n_g^2)(n_1^2+n_2^2),\\
N_{10} &= (n_1+n_2)^2(n_1-n_g)^2(n_2-n_a)^2,\\
N_{11} &= (n_1-n_2)^2(n_1+n_g)^2(n_2-n_a)^2,\\
N_{12} &= -2(n_1^2-n_2^2)(n_1^2-n_g^2)(n_2-n_a)^2,\\
N_{13} &= (n_1-n_2)^2(n_1-n_g)^2(n_2+n_a)^2,\\
N_{14} &= (n_1+n_2)^2(n_1+n_g)^2(n_2+n_a)^2,\\
N_{15} &= -2(n_1^2-n_2^2)(n_1^2-n_g^2)(n_2+n_a)^2,\\
N_{16} &= 2(n_1^2-n_2^2)(n_2^2-n_a^2)(n_1-n_g)^2,\\
N_{17} &= 2(n_1^2-n_2^2)(n_2^2-n_a^2)(n_1+n_g)^2,\\
N_{18} &= -4(n_2^2-n_a^2)(n_1^2-n_g^2)(n_1^2+n_2^2).
\end{aligned}$$

In terms of the above coefficients and the phase angles ϕ_1, ϕ_2, and ϕ_a, the denominator of Eq. (5) can be numerically evaluated with relative ease. The calculations are particularly simple for the case in which $\phi_1=\phi_2$. For this case, one may write

$$\begin{aligned}
\Delta\Delta^*=a&+b\cos 2\phi_1+c\cos 4\phi_1+d\cos 6\phi_1\\
&+e\cos 8\phi_1+f\cos(2\phi_1+2\phi_a)\\
&+g\cos(2\phi_1-2\phi_a)+h\cos(4\phi_1+2\phi_a)\\
&+j\cos(4\phi_1-2\phi_a)+k\cos(6\phi_1+2\phi_a)\\
&+1\cos(6\phi_1-2\phi_a)+m\cos(8\phi_1+2\phi_a)\\
&+n\cos(8\phi_1-2\phi_a)+p\cos 2\phi_a,
\end{aligned} \quad (6)$$

where

$$\begin{aligned}
a=\sum_{s=1}^{18}N_s^2+2(&N_2N_4+N_2N_9+N_3N_7\\
&+N_4N_9+N_6N_8+N_{11}N_{13}+N_{11}N_{18}\\
&+N_{12}N_{16}+N_{13}N_{18}+N_{15}N_{17}),
\end{aligned}$$

and

$$\begin{aligned}
b=2(&N_1N_3+N_1N_7+N_2N_3+N_2N_6+N_2N_7\\
&+N_2N_8+N_3N_4+N_3N_9+N_4N_6+N_4N_7\\
&+N_4N_8+N_5N_6+N_5N_8+N_6N_9+N_7N_9\\
&+N_8N_9+N_{10}N_{12}+N_{10}N_{16}+N_{11}N_{12}+N_{11}N_{16}\\
&+N_{11}N_{17}+N_{12}N_{13}+N_{12}N_{18}+N_{13}N_{15}\\
&+N_{13}N_{16}+N_{13}N_{17}+N_{14}N_{15}+N_{14}N_{17}\\
&+N_{15}N_{18}+N_{16}N_{18}+N_{17}N_{18}+N_{11}N_{15}).
\end{aligned}$$

Similar expressions for the other coefficients in Eq. (6) are readily obtained, but will not be reproduced here. Equation (4) is regarded as the fundamental formula in developing expressions convenient for numerical work.

3. A FABRY-PEROT INTERFEROMETER WITH HIGH TRANSMISSIVITY

If in addition to letting the optical thicknesses of the two layers be equal, one assumes that the coatings are "quarter-wave" layers, Eq. (6) becomes

$$\Delta\Delta^*=\alpha-\beta\cos 2\phi_a, \quad (7)$$

where

$$\alpha=a-b+c-d+e,$$

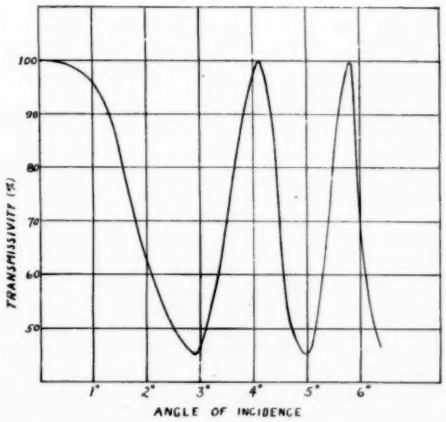


FIG. 2. Intensity distribution of transmitted radiation for small angles of incidence with optically treated interferometer. Ordinates are percent of incident intensity transmitted.

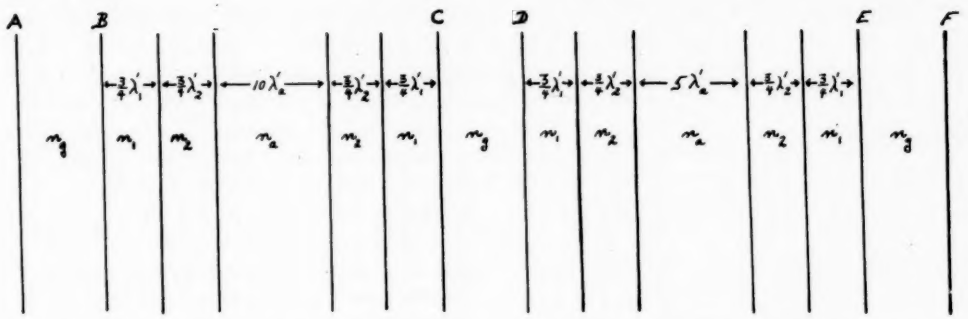


FIG. 3. Schematic diagram of simple two-component interference filter. Surfaces A and F are coated for zero reflectivity for the wave-length λ_a'.

and

$$\beta = f+g-h-j+k+l-m-n-p.$$

If we let

$$\gamma^2 = (64)^2 n_1^4 n_2^4 n_a^2 n_g^2,$$

we have for the transmissivity of the system

$$T = \gamma^2/(\alpha - \beta \cos 2\phi_a) = T_{\max}/(1+\delta \sin^2 \phi_a), \quad (8)$$

where

$$T_{\max} = \gamma^2/(\alpha-\beta),$$

and

$$\delta = 2\beta/(\alpha-\beta).$$

Equation (8) is in the standard form for the transmissivity of the Fabry-Perot interferometer for perpendicularly incident light. Now it will be observed that, for small angles of incidence, the phase shift resulting from the obliquity of the direction of propagation in the two optical coatings will be negligibly small compared to that incurred in traversing the air space, provided that the value of h_a is much greater than that of the coatings. It is easy to show that for small angles of incidence, θ, the transmissivity is given by the approximately accurate formula

$$T = \gamma^2/(\alpha - \beta \cos 2\phi_a \eta), \quad (9)$$

where

$$\eta = \sec \theta.$$

In Fig. 2 is plotted the transmissivity of such an interferometer as described in the preceding section plotted as a function of the angle of incidence of the radiation on the first glass surface. The following numerical data were used in the computations for Fig. 2:

$$n_1 = 1.40, \quad n_2 = 1.80, \quad n_g = 1.60, \quad n_a = 1.00,$$
$$h_1 = \lambda_a/4n_1, \quad h_2 = \lambda_a/4n_2, \quad h_a = 2\times 10^4 \lambda_a.$$

The distribution of intensity predicted in Fig. 2 is, of course, just that normally observed with the conventional silver-surfaced Fabry-Perot instrument. However, the "visibility," defined as $(T_{\max}-T_{\min})/(T_{\max}+T_{\min})$, is considerably less than can be obtained with silver-surfaced plates. In this example, the only advantage to be found in the high reflecting optical coating lies in the fact that the peaks of the transmissivity curve represent 100 percent transmission as opposed to the low values obtainable with the conventional instrument. The main disadvantages in the optically treated interferometer appear to be the relatively low visibility and the fact that the plates must be coated for just one wavelength—if quarter-wave layers are used.

4. AN EFFICIENT INTERFERENCE FILTER

The principle of an interference monochromator is well known. With the advent of high reflecting and practically non-absorbing optical coatings, it appears that the construction of an efficient interference filter of conveniently small size may be a practical possibility.

As a simple type of such an efficient monochromator, the device schematically diagrammed in Fig. 3 has been analyzed by means of the formulas developed in the preceding sections. In order to avoid the possibility of interference between the two units I and II, it is assumed that the set of planes labeled A, B, and C is inclined at a small angle (about one-half degree) to the set of planes labeled D, E, and F. Equation (5) predicts the behavior of each of the interferometers, I and II. If the transmissivities of the two units are represented by T_{I}

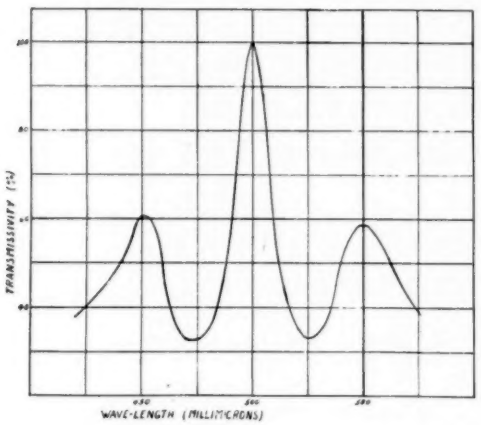


FIG. 4. Spectral distribution of transmitted radiation of simple interference filter designed for maximum at 500 millimicrons. Ordinates are percent of incident radiation transmitted. The ordinate identifications should be 450, 500, and 550.

and T_{II}, respectively, the transmissivity of the simple filter as diagrammed will be given by $T_I T_{II}$. To illustrate the performance of the above system, calculations of the transmissivity of the filter have been made with the following values of the constants involved: $n_1 = 1.40$, $n_2 = 1.80$, $n_g = 1.60$, $n_a = 1.00$, $\lambda_a' = n_1\lambda_1' = n_2\lambda_2' = 500$ mμ. The results of the calculation are shown graphically in Fig. 4.

A few well-known facts regarding the type of filter considered here are obvious from the derived formulas. The half-width of the transmitted band is smaller the larger the separation of the plates. When the order of interference is large for a given section of the filter, the transmissivity *versus* wave-length curve for that section exhibits maxima which are close together. For example, section I of the filter in Fig. 3 gives maximum transmissivities at about 452, 500, and 550 mμ. The maxima given by section II are spaced approximately twice as far apart. It appears that best results will be obtained by making the thicknesses of the optical coatings one-quarter wave-length instead of larger odd multiples of this thickness. The surface of the filter upon which the radiation is initially incident and the surface from which it is finally emergent should be coated for zero reflectivity for the wave-length corresponding to the peak of the transmissivity curve. The rate of development of heat in the filter will be small due to the fact that the radiation that is not transmitted is reflected instead of absorbed.

By the method outlined and demonstrated in this paper, it is possible to design a filter which will satisfy any reasonable demands involving the transmission of radiation over one limited part of the visible spectrum. It is not difficult to develop a general theory of the characteristics of multilayer optical coatings; however, it is believed that the direct approach demonstrated here will in general be of more value to the designer and computer. In any event, a considerable amount of numerical work is involved in determining transmissivities as functions of wave-length.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 36, NUMBER 5 MAY, 1946

The Effect of Wave-Length Distribution on the Brightness of Phosphors *

R. T. ELLICKSON
The Polytechnic Institute of Brooklyn, Brooklyn, New York
(Received February 28, 1946)

The visual efficiency of a phosphor may be defined as the visual response produced by the phosphor per quantum of emitted light divided by the visual response per quantum at the maximum of the visibility curve for the eye. A simple mathematical expression is derived for this quantity that is valid for phosphors with an emission curve that can be represented by a Gauss error function. Families of curves are given that enable one to read off the visual efficiency of a phosphor, given the peak wave-length and half-width of its emission curve. When the peak of the emission curve is separated from the peak of the visibility curve by more than a critical value of about 400A, there is a value of the half-width of the emission curve different from zero for which the visual efficiency is a maximum.

IN the preparation of phosphorescent and fluorescent materials it is often of interest to know the visual efficiency of the emitted light, i.e., the visual response produced by the phosphorescent light per emitted quantum. A simple graphical method for estimating this quantity is described herewith.

The apparent brightness of a phosphor is determined by three quantities: (1) the number of quanta emitted per unit area per second, (2) the wave-length distribution of the emitted light, and (3) the visibility curve of the eye. In mathematical form, the brightness is given by

$$B=K\int_0^\infty P(\lambda)\,V(\lambda)d\lambda, \qquad (1)$$

where $P(\lambda)$ is a function representing the emission of the phosphor, $V(\lambda)$ is a function representing the relative response of the eye to different wave-lengths, and K is a constant proportional to the number of quanta emitted per second by the phosphor.

In general, the functions $P(\lambda)$ and $V(\lambda)$ are not given in analytic form, and the integration of Eq. (1) is carried out by multiplying the two curves point by point and integrating the resulting curve by graphical methods. It happens that the emission curves of many phosphors can be reasonably well represented by a Gauss error function, and this is also true for the visibility curve of the eye. If such a function is used to represent these quantities, Eq. (1) can be readily integrated. The resulting equation for the relative brightness is very simple in form, and is quite useful for a quick estimation of the effect of the wave-length distribution of a particular phosphor on its brightness.

The visibility curve of the eye is quite well represented by the equation

$$V(\lambda)=\exp[-\alpha(\lambda-\lambda_1)^2],$$

where

$\alpha=3.75\times10^{-6}$, $\lambda_1=5100$A, for scotopic vision,

and

$\alpha=2.77\times10^{-6}$, $\lambda_1=5550$A, for photopic vision.

Suppose the emission curve for the phosphor is given by

$$P(\lambda)=\exp[-\beta(\lambda-\lambda_2)^2].$$

Then the brightness of the phosphor will be

$$B=K\int_0^\infty \exp\,[-\alpha(\lambda-\lambda_1)^2-\beta(\lambda-\lambda_2)^2]d\lambda,$$
$$=K\left[\frac{\pi}{\alpha+\beta}\right]^{\frac{1}{2}}\exp\left[-\frac{\alpha\beta}{\alpha+\beta}(\lambda_2-\lambda_1)^2\right]. \qquad (2)$$

The area under the phosphor curve itself is given by

$$A=K\int_0^\infty \exp\,[-\beta(\lambda-\lambda_2)^2]d\lambda,$$
$$=[\pi/\beta]^{\frac{1}{2}}. \qquad (3)$$

* The work described in this paper was carried out, in whole or in part, under Contract OEMsr-982 between the Office of Research and Development and the Polytechnic Institute of Brooklyn.

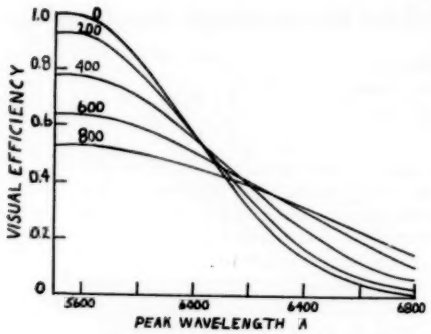


FIG. 1. Visual efficiency for photopic vision for various half-widths.

The relative efficiency of the phosphor is best represented by dividing the brightness by this area. This has the effect of giving a visual efficiency for the phosphor. A phosphor emitting all its light at the peak of the visibility curve would thus have a visual efficiency of unity. Dividing Eq. (2) by Eq. (3), the visual efficiency of the phosphor is

$$E=\left[\frac{\beta}{\alpha+\beta}\right]^{\frac{1}{2}}\exp\left[-\frac{\alpha\beta}{\alpha+\beta}(\lambda_2-\lambda_1)^2\right].$$

The quantity, α, is related to the half-width at half-maximum, $\Delta\lambda$, by the equation

$$\alpha=0.693/(\Delta\lambda)^2.$$

Hence

$$E=\left[1+\left(\frac{\Delta\lambda_2}{\Delta\lambda_1}\right)^2\right]^{-\frac{1}{2}}\exp\left[-\frac{0.693}{\Delta\lambda_1{}^2+\Delta\lambda_2{}^2}(\lambda_2-\lambda_1)^2\right].$$

For photopic vision, this becomes

$$E=\left[1+\left(\frac{\Delta\lambda}{500}\right)^2\right]^{-\frac{1}{2}}\times\exp\left[-\frac{2.77\times10^{-6}}{1+(\Delta\lambda/500)^2}(\lambda-5550)^2\right], \quad (4)$$

and for scotopic vision

$$E=\left[1+\left(\frac{\Delta\lambda}{430}\right)^2\right]^{-\frac{1}{2}}\times\exp\left[-\frac{3.75\times10^{-6}}{1+(\Delta\lambda/430)^2}(\lambda-5100)^2\right], \quad (5)$$

where λ and $\Delta\lambda$ are the peak wave-length and half-width at half-maximum for the phosphor.

An example will show how this works for a particular sample. Strontium selenide[1] activated with samarium and europium has its peak emission at 5750A and a half-width of 250A. Substituting these values in Eq. (5) gives a value for E of 0.26. Measurement of the curve obtained by point by point multiplication of the visibility and emission curves gives a value of 0.24, which is a reasonably good check. As another example, strontium sulfide[2] activated by samarium and europium has a peak wave-length of 6350A and a half-width of 400A. Substitution of these values in Eqs. (4) and (5) gives values of 0.26 and 0.030 for the photopic and scotopic visual efficiencies respectively, which compares well with the results of the direct measurement with a planimeter, 0.24 and 0.026.

These results can be put in a convenient graphical form. Figures 1 and 2 show the variation of visual efficiency with the position of the peak of the emission curve and the half-width for photopic vision, and Figs. 3 and 4 show the same thing for scotopic vision. The visual efficiency is symmetrical about the values 5550 and 5100 in Figs. 1 and 3, and therefore only the higher wave-length portion of the curves is shown. For samples that have their peaks below these values the same curves can be used. Thus a sample with its peak emission at 5200A would have the same visual efficiency as one peaking at 5900A for photopic vision and 5300A for scotopic vision.

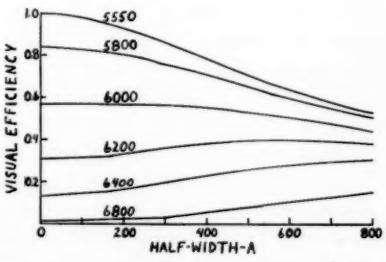


FIG. 2. Visual efficiency for photopic vision for various peak wave-lengths.

[1] The properties of this phosphor were described by R. Ward at the March, 1946 meeting of the Optical Society of America.

[2] See report of F. Urbach at the same meeting.

It is interesting that for phosphors that have their peak emission near the peak of the visibility curve the visual efficiency decreases with increasing half-width, while the reverse is true for those where the peaks are more widely separated. At an intermediate value of around 400A for the separation between the peaks, the value of the half-width over a wide range of values has little effect on the visual efficiency. This can be seen readily by differentiating Eq. (4) and setting the derivative equal to zero. Doing this, and solving for $\Delta\lambda$, the half-width, one obtains

$$\Delta\lambda = 500[5.54\times10^{-6}(\lambda-5550)^2-1]^{\frac{1}{2}}. \quad (6)$$

This indicates that there will be a value of $\Delta\lambda$ other than zero for which the visual efficiency will be a maximum whenever the peak of the emission differs from 5550A by more than 425A. The corresponding expression from Eq. (5) is

$$\Delta\lambda = 430[7.5\times10^{-6}(\lambda-5100)^2-1]^{\frac{1}{2}}, \quad (7)$$

indicating a minimum separation of the peaks of 365A in order for a maximum to occur. As an example, substituting $\lambda = 6200$A in Eq. (6) one obtains $\Delta\lambda = 580$A as the half-width for maximum visual efficiency. This value seems reasonable from a glance at Fig. 2.

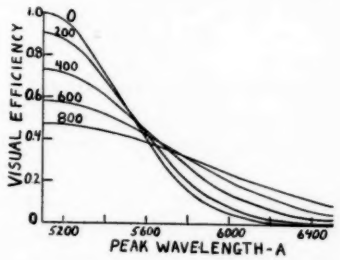


FIG. 3. Visual efficiency for scotopic vision for various half-widths.

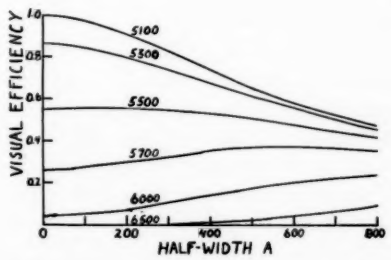


Fig. 4. Visual efficiency for scotopic vision for various peak wave-lengths.

The half-width of the emission curve of a phosphor is usually changed considerably by cooling it to a low temperature. Evidently, aside from effects on the quantum efficiency of the phosphor, this could change the brightness in either direction, depending on the position of the peak of the emission curve. For instance, the strontium sulfide phosphor mentioned above, which has a visual efficiency for photopic vision of 0.26 at room temperature would have this reduced to 0.17 if cooled sufficiently to change its half-width to 200A. Its visual efficiency for scotopic vision would drop from 0.030 to 0.008, or by nearly a factor of four. On the other hand, as a glance at Fig. 3 shows, a phosphor with its peak emission at 5600A would change very little in brightness with drastic changes in its half-width.

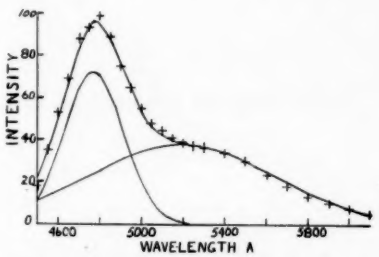


FIG. 5. Emission of strontium sulfide phosphor under infra-red stimulation. The upper curve is the sum of the two Gauss error curves shown.

Of course there are many phosphors whose emission curves would not be well represented by a Gauss error function. For those that have a complicated shape one would have to revert to the method of point by point multiplication. There are some, however, that can be well represented by two or more error functions, with different peaks and half-widths. For these one could use the present method, determining the efficiencies of each component separately. The over-all efficiency would be the weighted average of the separate efficiencies.

The emission curve for such a sample is shown in Fig. 5, which is for a strontium sulfide phosphor activated with cerium and samarium.[2] The

emission curve is quite well represented by

$$B = 72 \exp(-2.55\times10^{-5}(\lambda-4770)^2), \\ +38 \exp(-2.57\times10^{-6}(\lambda-5200)^2).$$

The half-widths are 165A and 520A, respectively. Substituting these values in the appropriate equations, one gets .652 and .625 for the scotopic, and .208 and .506 for the photopic efficiencies.

The area under the curve A exp $(-\alpha x^2)$ is equal to $A(\pi/\alpha)^{\frac{1}{2}}$. The areas under the two curves are therefore in the ratio 1.66:1. The weighted average gives .63 and .39 for the scotopic and photopic efficiencies of the phosphor.

The author is indebted to Dr. R. Ward under whose supervision the phosphor samples were made, and to Dr. W. L. Parker who determined the shape of the emission curves.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 36, NUMBER 5 MAY, 1946

Light Sum of Phosphors under Thermal and Infra-Red Stimulation *

R. T. ELLICKSON
The Polytechnic Institute of Brooklyn, Brooklyn, New York
(Received February 28, 1946)

The experiments reported deal with the recently discovered type of phosphor that shows a large amount of luminescence of visible light when irradiated with infra-red. The emission can be made to occur when the normal phosphorescence has decayed to a negligible value. The light sum, or area under the decay curve, of such phosphors under infra-red stimulation is of the order of 1000 microlambert minutes. Measurements are reported for the light sum when the sample is irradiated with infra-red and when it is heated. In some phosphors the infra-red light sum is from 15 to 30 times as great as the light sum obtained by heating. In others, it is about the same, or perhaps even less. Attention is called to the connection between these experiments and the "glow curve" experiments used to determine the distribution of electron traps in phosphors.

INTRODUCTION

THE rate of decay of the light emission from a phosphor can be changed by heating, and sometimes by irradiation with infra-red light, hereinafter denoted by *IR*. The usual effect of the *IR* is to increase the brightness of the phosphor very slightly. It sometimes has the opposite effect in the phenomenon known as quenching. There have recently been described a group of phosphors[1] that show a very marked increase in their brightness upon irradiation with *IR*, even when the natural phosphorescent brightness has dropped to a negligible value. It is with this type of phosphor that this report is concerned.

The release of electrons from traps by heating has been used extensively to study the distribution of electrons in traps of various depths in phosphors. The phosphor is excited at a low temperature, usually by irradiation with ultraviolet light, and the light emitted by the phosphor is measured as the temperature is raised at a known rate. Quite an extensive theory has been developed to relate the shape of these "glow curves" to the electron trap distribution in the phosphor.[2]

It is the purpose of this paper to show that the ratio of the amount of light emitted by irradiating the phosphor with *IR* to that emitted by heating the phosphor above room temperature is by no means constant for different phosphors.[3] In some cases the ratio seems to be as large as 30, and in others there are indications

* The work described in this paper was carried out, in whole or in part, under Contract OEMsr-982 between the Office of Scientific Research and Development and the Polytechnic Institute of Brooklyn.

[1] See reports of F. Urbach, G. R. Fonda, R. Ward, *et al.* at the March, 1946 meeting of the Optical Society of America.

[2] J. T. Randall and M. F. H. Wilkins, Proc. Roy. Soc. **184**, 366 (1945).

[3] This fact was first called to my attention by F. Urbach.

that the ratio may be less than one. This would indicate that the probability for non-radiative transitions by the trapped electrons is a marked function of the temperature. It would seem, therefore, that some account of this fact should be taken in interpreting the results of glow curve experiments.

APPARATUS

The experimental arrangement is shown schematically in Fig. 1. The samples were round buttons 1.25 inches in diameter and about 0.13 inch thick, with one surface flat and one slightly rounded. The light emitted by the phosphor was picked up by an electron multiplier tube, RCA 1-P-22. The potential difference between stages was $112\frac{1}{2}$ volts, furnished by a battery power supply, making the amplification essentially constant over relatively long periods of time. The current in the final stage, without further amplification, was read on a galvanometer equipped with an Ayrton shunt. The sensitivity of the galvanometer as set up was 3×10^{-10} amp. per mm. The maximum deflections observed in these experiments were about 500,000 divisions. Fluctuations in the dark current of the cell, which was about 200 divisions, were quite negligible.

The heater coil was in a porcelain holder in the form of a spiral, the diameter of the coil being about an inch. Provision was made for varying the current in the heating coil.

The *IR* was from a 10-watt tungsten lamp placed 18 inches from the sample and in a position such that the light fell on the sample at nearly normal incidence. The light from this lamp was filtered through a 5-mm Corning No. 2540 filter.

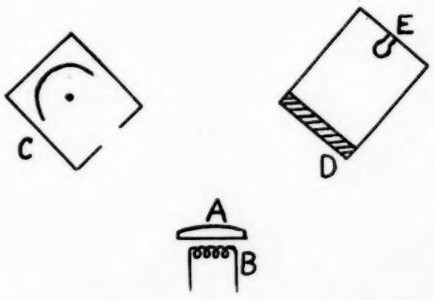


FIG. 1. Schematic diagram of apparatus. *A*—Phosphor sample, *B*—spiral heating coil, *C*—RCA 1-P-22 electron multiplier tube, *D*—Corning 2540 filter, *E*—tungsten lamp.

EXPERIMENTAL RESULTS

The sample was fully excited, placed on the heater with the current shut off, and the *IR* was turned on. Readings of the current from the electron multiplier tube were taken every 10 seconds for the first few minutes, and at more widely spaced intervals thereafter. The sample was then re-excited, placed on the heater in an identical position, the heater was switched on, and the current again read every 10 seconds until the reading had passed through a maximum and reached zero. To insure that no light was left in the phosphor, the *IR* was turned on at the end of the heating period. In all cases the *IR* produced no perceptible deflection, indicating that the phosphor was completely exhausted by the heating process.

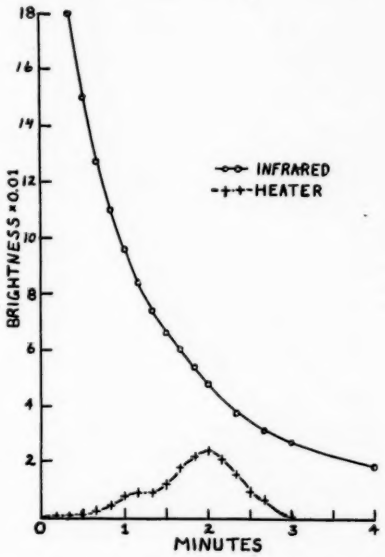


FIG. 2. Exhaustion of Sr S-Ce-Sm phosphor under infra-red and thermal stimulation.

The decay of a typical sample of strontium sulfide[4] activated with cerium and samarium is shown in Fig. 2. The light sum, which is propor-

[4] The fact that this phosphor shows a marked response to infra-red light was discovered by Dr. F. Urbach, of the University of Rochester.

tional to the area under the curve, is obviously much larger when the sample is irradiated with *IR* than when it is heated. In Fig. 3 the reciprocal of the square root of the brightness is plotted against the time. The approximate linearity of the points indicates that the data can be well represented by an equation of the form

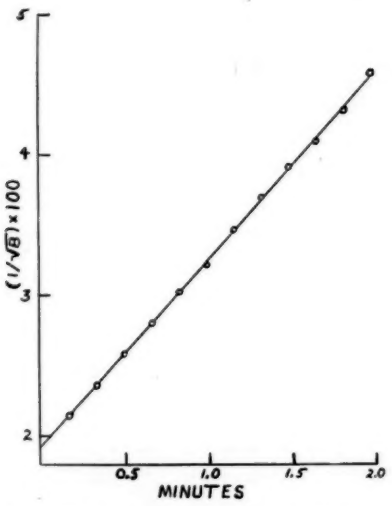


FIG. 3. Sr S-Ce-Sm phosphor. The straight line indicates the validity of Eq. (1).

$$B = Bo(1+at)^{-2}. \quad (1)$$

The light sum, which is the integral of this expression, is given by

$$L = Bo/a.$$

For these data, Bo and a have the values 2710 and 0.683, from which the light sum, L, is equal to 3960 division minutes. With a Macbeth illuminometer, this is found to be equivalent to 1200 microlambert minutes.

The data for this run is given in Table I. Here the light sum is also calculated, for the first 40 minutes of exhaustion, by averaging each pair of adjacent points and multiplying by the length of the interval. The value of the light sum for *IR* irradiation obtained in this way, 4110, compares favorably with the value 3960 obtained by integrating the equation of the curve. One would expect to find the actual light sum to be larger than that given by integrating Eq. (1), because of the effect of electrons lying in the deeper traps. Experiments to be described in a different communication indicate that the actual light sum is about two or three times that given by integration of the equation describing the brightness in the first few minutes of decay.

The light sum obtained by heating, 290, is about seven per cent of that obtained by *IR* irradiation of this sample as determined here. The actual ratio, if the samples were taken to complete exhaustion, would doubtless be even smaller, since the samples still showed some response to *IR* at the end of the *IR* run, but none at all at the end of a heating run.

Similar experiments with a strontium selenide[5] phosphor activated with samarium and europium are shown in Fig. 4. Again it is obvious that the light sum under *IR* irradiation is far greater than with the heater. The plot in Fig. 5 of the reciprocal of the square root of the brightness against time indicates that for this sample also the decay of the brightness is well represented by Eq. (1) in the early stages of decay. From other work it is known that at the later stages of

TABLE I. Exhaustion of Sr S-Ce-Sm phosphor, 5-minute ultraviolet excitation.

Time	Infra-red irradiation Brightness	$B\Delta t$	Heater Brightness	$B\Delta t$
0	2780*	—	0	—
10″	2190	414	4.5	—
20″	1800	332	6	1
30″	1500	275	10	1
40″	1280	232	21	3
50″	1100	198	45	6
1′	960	172	73	10
10″	840	150	88	13
20″	740	132	90	15
30″	660	117	120	18
40″	600	105	180	25
50″	540	92	220	33
2′	480	73	240	38
10″	—	—	210	38
20″	375	146	155	30
30″	—	—	95	21
40″	315	115	65	13
3′	272	98	9.5	12
4′	188	230	2.5	6
5′	138	163	1.5	2
10′	51	470	0	4
15′	28	197	0	0
20′	19	117		0
40′	10	120		0
Total		4113		289

* Extrapolated value.

[5] The large response of this phosphor to infra-red illumination was discovered by Dr. R. Ward of this institute.

the decay, the brightness is higher than given by the equation, but integration of the equation, with values of *Bo* and *a* of 233 and 0.139 from Fig. 5, will at least give a lower limit to the light sum. This gives value of 1680 division minutes, equivalent to 1250 microlambert minutes. Table II shows the data for the sample under *IR* irradiation, and heated with three different values of current. The light sum seems to be smaller by at least a factor of ten when the sample is heated. The light sum seems to be different for different heating rates, with a maximum for an intermediate rate. This may be connected with the phenomenon described by Kroger,[6] who found that a ZnS-Mn phosphor showed a maximum luminescence at about 80°C.

A SrS[7] phosphor activated by samarium and europium behaves quite differently from the two phosphors described above. For one thing, the plot of the reciprocal of the square root of the brightness against time, shown in Fig. 6, is not a straight line. It is therefore not easy to obtain the light sum by integration, since the equation of the decay curve is not readily apparent. A somewhat different procedure was therefore adopted for determining the light sum in this case.

The 10-watt tungsten lamp was replaced by two 100-watt lamps, which were connected to a variac. When the switch was turned on, the voltage across the lamp was gradually increased by manual adjustment, so as to keep the light output of the phospher constant. This was repeated at three different levels of light intensity as shown in Fig. 7.

The end of the flat portion of curve indicates the time at which the variac was turned on fully,

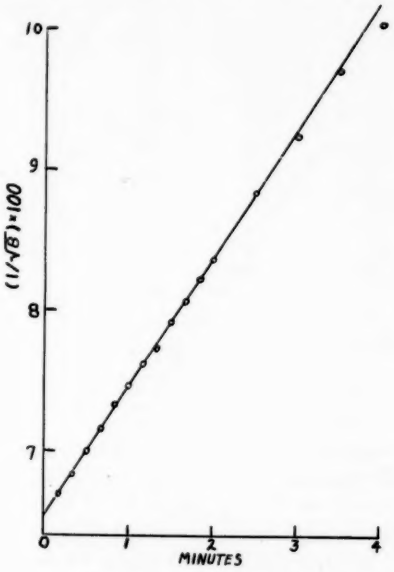


FIG. 5. Sr Se-Sm-Eu phosphor.

and the rate of decay down to 100 divisions was noted. The "glow curve" is shown on the same graph. In this case, the light sum, measured to the point that the brightness falls to 100 divisions, from heating is greater than from *IR* irradiation by more than a factor of 3. In the first case it is 4000 division minutes, equivalent to 2600 microlambert minutes. In the second, the average of the three runs is about 1200 division minutes. The ratio should probably be smaller for two reasons; (1) the rate of decay from *IR* illumination is probably smaller than for the heated sample beyond the point where readings were taken; (2) upon heating, the color of this sample changed from red to yellow, which would have a greater response from the photocell used. It is not unreasonable that, if these two effects were corrected for, the light sum would be about the same for this sample for either mode of stimulation.

A repeat run on the Sr S-Ce-Sm sample was made using the method just described. The re-

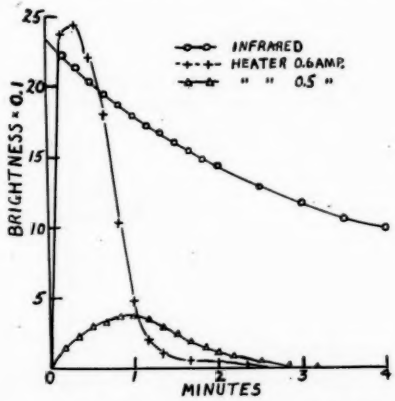


FIG. 4. Exhaustion of Sr Se-Sm-Eu under infra-red and thermal stimulation.

[6] F. A. Kroger, Physica 6, 369 (1939).
[7] Also discovered by Dr. Urbach.

TABLE II. Exhaustion of Sr Se-Sm-Eu phosphor, 5-minute excitation with blue light.

Time	Infra-red irradiation Brightness	$B\Delta t$	0.5 amp. B	$B\Delta t$	0.6 amp. B	$B\Delta t$	0.7 amp. B	$B\Delta t$
0	232*	—	0	—	0	—	0	—
10″	222	37.8	14	1.2	237	19.8	462	38.5
20″	214	36.1	22	3.0	244	40.1	237	58.2
30″	204	34.8	29	4.3	220	38.6	41	23.2
40″	195	33.2	32	5.1	180	33.3	9	4.2
50″	187	31.8	36	5.7	103	23.6	5	1.2
1′	179	30.5	36	6.0	48	12.6	4	0.8
10″	172	29.2	34	5.8	20	5.7	4	0.7
20″	167	28.2	29	5.3	10	2.5	2	0.5
30″	160	27.2	24	4.4	7	1.4	1	0.3
40″	154	26.2	18	3.5	5	1.0	0	0.1
50″	148	25.2	14	2.7	4	0.8	0	0
2′	143	24.2	10	2.0	4	0.7	0	0
30″	128	67.7	3	3.3	2	1.5	0	0
3′	117	61.3	1	1.0	1	0.8	0	0
30″	106	55.8	0	0.3	0	0.3	0	0
4′	99	51.3	0	0	0	0	0	0
30″	91	47.5	0	0	0	0	0	0
5′	83	43.5	0	0	0	0	0	0
30″	76	39.8	0	0	0	0	0	0
Total		731		54.6		183		128

* Extrapolated value.

sults are shown in Fig. 8. The light sum from *IR* stimulation, taken only to a brightness of 100, is over five times as great with *IR* stimulation as with heating. There was no noticeable color change in this sample on heating. To determine the validity of the method used in the earlier part of the paper for determining the light sum, the decay of this sample was measured at three different intensities. The results are plotted in Fig. 9. The straight lines obtained indicate the applicability of the equation

$$B = Bo(1+at)^{-2}$$

at least in the early stages of the decay. The values of Bo, a, and the light sum $L = Bo/a$ are shown in Table III. The average light sum here, 1150, is about two-thirds that obtained from measuring the area under the curve in Fig. 8. Thus, one is certainly conservative in using Bo/a as the light-sum under *IR* stimulation.

DISCUSSION

It is to be expected that, as the temperature is raised, the probability that the energy of the excited electron in a phosphor be released in the form of heat, and therefore not in the form of radiation, is increased.[8] This would account for

[8] Mott and Gurney, *Electronic Processes in Ionic Crystals* (Oxford University Press, New York), p. 219.

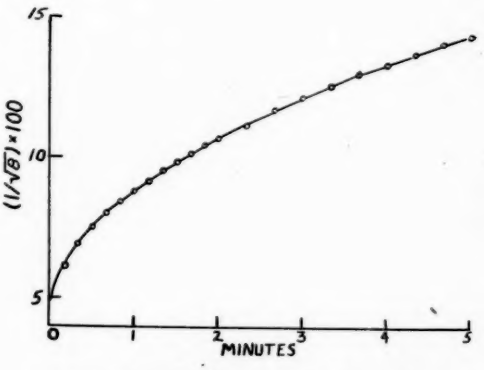


FIG. 6. Sr S-Sm-Eu phosphor. The decrease in slope can be attributed to the emptying of shallow traps.

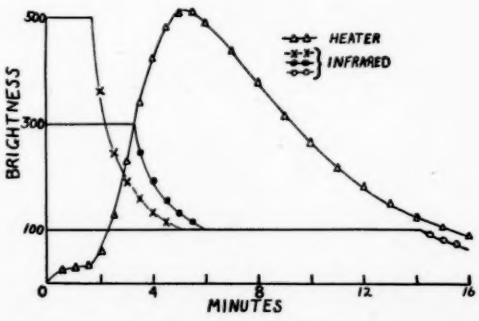


FIG. 7. Exhaustion of Sr S-Sm-Eu.

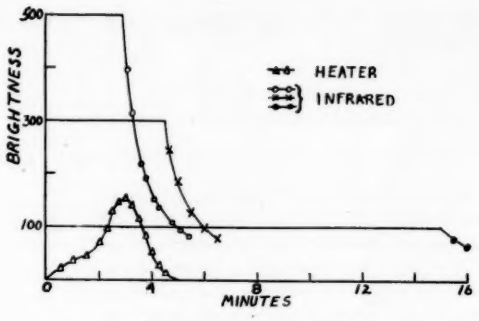


FIG. 8. Exhaustion of Sr S-Ce-Sm.

the behavior of those phosphors that show a much smaller light sum under thermal stimulation than under *IR* stimulation. Since the effect is so pronounced, it would seem essential that it be taken into account in using "glow curves" to obtain electron trap distribution in phosphors. It seems likely that the introduction of this factor, which would not be easy to do, would not have too much of an effect on the position of the traps on an energy scale, but it would have a decided effect on the relative population of electrons in those traps.

TABLE III.

	Bo	*a*	*L* = *Bo*/*a*
1	202	.185	1100
2	465	.41	1130
3	935	.76	1230

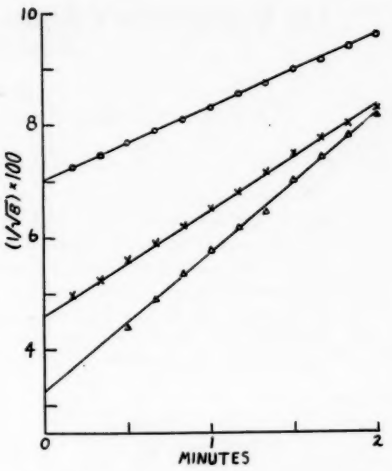


FIG. 9. Exhaustion of Sr S-Ce-Sm under three different levels of IR intensity.

The author is indebted to Mr. W. Primak for preparing the samples used here and to Dr. W. L. Parker who built the electron multiplier apparatus.

The Evaluation of Colorimetric Tristimulus Values by Means of a New Centroid Method *

HOWARD A. ROBINSON
Research Laboratories, Armstrong Cork Company, Lancaster, Pennsylvania.

(Received March 8, 1946)

A new semigraphical method of integration which is peculiarly suited to the evaluation of certain integrals which appear with great frequency in color specification work is introduced and discussed. Previous integration methods either require the use of special equipment or are extremely laborious and time consuming especially if the thirty ordinate or 100 ordinate methods are used. The present method usually reduces the number of ordinates to 8 and allows the use of either 4, 2, or even one ordinate in certain calculations. For essentially linear reflectance curves the tristimulus values can be read from the spectrophotometric curve directly. The saving in time is marked: The accuracy is sufficient for most industrial purposes.

INTRODUCTION

IN color measuring problems it is common practice to determine either the trichromatic coefficients or the tristimulus values of the color. The latter are defined by the integrals

$$X = \int E_L(\lambda) R(\lambda) \bar{x}(\lambda) d\lambda,$$

$$Y = \int E_L(\lambda) R(\lambda) \bar{y}(\lambda) d\lambda,$$

$$Z = \int E_L(\lambda) R(\lambda) \bar{z}(\lambda) d\lambda,$$

where $E_L(\lambda)$ is the relative energy function of the light source, as a function of the wave-length (λ), $R(\lambda)$ is the reflecting power of the pigment, and $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ are, respectively, the tristimulus values of the spectrum.

The values for E_L, $\bar{x}$, $\bar{y}$, $\bar{z}$ are well known and have been tabulated[1] together with the product functions $E_L\bar{x}$, $E_L\bar{y}$, and $E_L\bar{z}$. These latter can be conveniently denoted by ϕ_x, ϕ_y, and ϕ_z. Values of these ϕ_k will, of course, be different for each light source (illuminant A, B, or C) used.

Mathematically the problem reduces to the evaluation of integrals such as

$$\int_0^\infty R(\lambda)\phi_k d\lambda,$$

where ϕ_k is the tabulated product function and $R(\lambda)$ is an experimentally determined curve of reflecting power *vs.* wave-length. Such integrations may be conveniently carried out by a recently described method.[2]

The method requires the setting up of certain diagrams called scales which can be constructed from the ϕ_k product function alone. This is done once and for all. The scales are then applied to the spectrophotometric curves and the final evaluation follows graphically. Scales have been constructed for illuminants A, B, and C and will be found in this paper.

PART I. THEORY OF THE METHOD

It is convenient to consider a mechanical analogy as an aid in understanding the evalua-

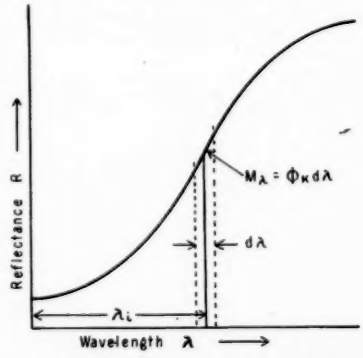


FIG. 1. The "moment of mass distribution" about the R axis will be given by $\int_0^\infty \lambda_i \phi_k d\lambda$ and about the λ axis by $\int_0^\infty R\phi_k d\lambda$.

* Presented at the Winter Meeting of the Optical Society of America, March 7–9, 1946, Clevelend, Ohio.

[1] Hardy, *Handbook of Colorimetry* (Technology Press, Cambridge, Massachusetts, 1936).

[2] Popoff, Quart. App. Math. **3**, 166 (1945); (Robinson, Quart. App. Math. **3**, 383 (1945)).

tion of the colorimetric integrals. Thus in the integral

$$X = \int_0^\infty R\phi_x d\lambda,$$

ϕ_x can be considered as a variable mass distribution, i.e., (mass/unit length) along the curve $R=f(\lambda)$. The centroid of this mass distribution will have the coordinates $(\bar{\lambda}_x, \bar{R}_x)$. These may be found as follows: each element of length $d\lambda$ along the curve R will have a mass $M_\lambda = \phi_k d\lambda$. Thus about the R axis (see Fig. 1) the moment of mass distribution will be given by

$$\sum_{\lambda=0}^{\lambda=\infty} M_\lambda \lambda_i = \int_0^\infty \phi_k \lambda d\lambda,$$

and about the λ axis the moment of mass distribution will be given by

$$\sum_{\lambda=0}^{\lambda=\infty} R_i M_\lambda = \int_0^\infty R\phi_k d\lambda.$$

Since the total mass Ω_k is given by

$$\int_0^\infty \phi_k d\lambda$$

the centroid coordinates can be found from the equations

$$\bar{\lambda}_k = \int_0^\infty \lambda\phi_k d\lambda \Big/ \Omega_k, \qquad (1)$$

$$\bar{R}_k = \int_0^\infty R\phi_k d\lambda \Big/ \Omega_k. \qquad (2)$$

In colorimetric terms Ω_x is the area under the curve $E_L\bar{x}=f(\lambda)$ that is the integrated product of the light intensity function and the $\bar{x}$ luminosity function over the entire wave-length range. It will be noticed that the centroid abscissa $\bar{\lambda}_x$ is independent of the reflectance curve, and that the centroid ordinate $\bar{R}_x$, when multiplied by the area Ω_x, is the integral to be evaluated. $\bar{R}_x$ can be found graphically; Ω_x (and $\bar{\lambda}_x$) since they only depend on the product function can be determined once and for all. The evaluation of the colorimetric integral then follows from (2).

The graphical evaluation of $\bar{R}_k$ is accomplished by the use of two lemmas, proofs for which will be found in the appendix. Lemma 1 is restrictive

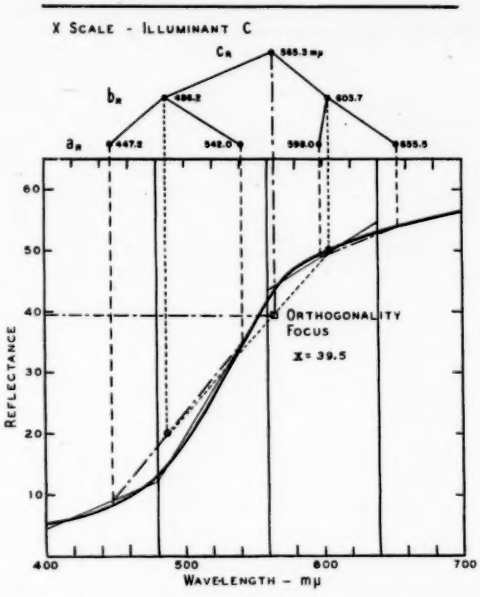


FIG. 2. Schematic spectrophotometric curve with illuminant C, x scale applied.

in the sense that it applies only to linear functions; it is thus necessary to replace the reflectance curve by a series of closely approximating linear segments. A certain amount of experience helps in carrying this out.

It is convenient to have a name for such mass distribution centroids determined as previously described: for this purpose the term orthogonality focus* (O.F.) was originally proposed. The term may be defined for any curve as the point which has the coordinates $(\bar{\lambda}_k, \bar{R}_k)$, when $\bar{\lambda}_k$ and $\bar{R}_k$ are given by Eqs. (1) and (2). Using these definitions it is possible to set up the necessary lemmas as follows:

Lemma 1. The orthogonality focus of a linear function lies on the straight line; thus $\bar{R}_k$ follows immediately once $\bar{\lambda}_k$ has been determined.

Lemma 2. If any wave-length interval of the reflectance curve be divided into two parts, the orthogonality foci of the two parts and of the whole are collinear.

* For details on reasons for selecting this term reference 1 may be consulted. The concept of orthogonality arises because of the theoretically possible relation $\bar{R}_k = \frac{1}{\Omega_x}\int_0^\infty R\phi_x d(\lambda) = 0$. In the colorimetric case this could only happen for a pure black ($R(\lambda)=0$) and hence is of little interest.

TABLE I. Comparison of three methods.

	Wt. ord.	Mech. integrator	O.F.	
Neutral photographic filter				
x	.3766	.373	.374	Read direct from curve
y	.3681	.371	.371	(1 abscissa)
Y	10.16%	10.4%	10.3%	
Natural cork				
x	.3929	.388	.387	2 abscissae used
y	.3705	.369	.370	
Y	34.8%	34.9%	35.4%	
Signal blue filter				
x	.2124	.211	.212	4 abscissae used
y	.2006	.200	.210	
Y	32.1%	31.9%	34.3%	
			.216	12 abscissae used
			.198	
			31.8%	
White paper*				
x	.3201	.320	.320	4 abscissae used
y	.3290	.329	.329	
Y	84.7%	84.6%	84.8%	

* Example given in Hardy, *Handbook of Colorimetry*.

The proofs of these lemmas will be found in the appendix. In some cases the term "mass distribution centroid" may be preferred to "orthogonality focus." An example of the application of Eqs. (1) and (2) will now be given.

In Fig. 2 a schematic spectrophotometric curve is reproduced. The curve has been arbitrarily split into four segments 400–480, 480–560, 560–640, and 640–700 $m\mu$ and the best linear approximation to each segment is drawn

TABLE II A. X Scale for A—Illuminant 400–700 Mu.

λ	X_1	X_2	X_3	X_4	X_5
400	414.33				
420		429.83			
421	432.25				
440			449.96		
441	450.64				
460		458.44			
461	469.06				
480				514.83	
481	486.79				
500		499.02			
501	515.90				
520			544.50		
521	532.95				
540		547.10			
541	551.89				
560					595.22
561	571.42				
580		583.12			
581	590.93				
600			601.63		
601	610.30				
620		618.33			
621	629.58				
640				608.94	
641	649.00				
660		654.03			
661	668.56				
680			656.35		
681	688.29				
700		688.29			

TABLE II B. Y Scale for A—Illuminant 400–700 Mu.

λ	Y_1	Y_2	Y_3	Y_4	Y_5
400	414.44				
420		431.85			
421	433.47				
440			465.26		
441	452.47				
460		467.42			
461	472.27				
480				532.52	
481	492.24				
500		506.95			
501	512.08				
520			534.45		
521	531.24				
540		542.48			
541	550.81				
560					572.07
561	570.49				
580		579.81			
581	590.16				
600			593.17		
601	609.84				
620		617.06			
621	629.37				
640				598.07	
641	648.92				
660		653.83			
661	668.53				
680			656.09		
681	688.28				
700		688.28			

TABLE II C. Z Scale for A—Illuminant 400–700 Mu.

λ	Z_1	Z_2	Z_3	Z_4	Z_5
400	414.34				
420		429.95			
421	432.32				
440			452.08		
441	450.87				
460		459.84			
461	469.66				
480				461.83	
481	489.04				
500		494.74			
501	508.77				
520			498.73		
521	528.64				
540		532.64			
541	548.05				
560					462.11
561	569.23				
580		576.74			
581	589.55				
600			582.30		
601	608.39				
620		612.31			
621	627.70				
640				582.46	
641	644.17				
660		644.17			
661	0				
680			644.17		
681	0				
700		0			

TABLE III B. Y Scale for B—Illuminant 400–700 Mu.

λ	Y_1	Y_2	Y_3	Y_4	Y_5
400	414.50				
420		431.72			
421	433.37				
440			463.74		
441	452.17				
460		466.46			
461	472.00				
480				528.45	
481	491.87				
500		505.55			
501	511.70				
520			531.95		
521	531.10				
540		541.97			
541	550.66				
560					561.45
561	570.24				
580		578.80			
581	589.91				
600			590.42		
601	609.70				
620		616.59			
621	629.25				
640				594.28	
641	648.83				
660		653.49			
661	668.40				
680			655.46		
681	688.11				
700		688.11			

TABLE III A. X Scale for B—Illuminant 400–700 Mu.

λ	X_1	X_2	X_3	X_4	X_5
400	414.40				
420		429.71			
421	432.14				
440			447.98		
441	450.33				
460		457.40			
461	468.81				
480				494.44	
481	486.55				
500		496.34			
501	515.70				
520			542.78		
521	532.84				
540		546.69			
541	551.73				
560					576.20
561	571.17				
580		582.14			
581	590.68				
600			599.04		
601	610.16				
620		617.85			
621	629.46				
640				605.19	
641	648.91				
660		653.68			
661	668.43				
680			655.71		
681	688.12				
700		688.12			

TABLE III C. Z Scale for B—Illuminant 400–700 Mu.

λ	Z_1	Z_2	Z_3	Z_4	Z_5
400	414.42				
420		429.83			
421	432.21				
440			449.99		
441	450.55				
460		458.75			
461	469.40				
480				456.78	
481	488.68				
500		493.43			
501	508.39				
520			496.33		
521	528.51				
540		532.27			
541	547.92				
560					456.89
561	568.98				
580		575.76			
581	589.32				
600			580.37		
601	608.28				
620		611.95			
621	627.65				
640				580.49	
641	644.07				
660		644.07			
661	0				
680			644.07		
681	0				
700		0			

in. The O.F. of each segment, which by lemma one lies on the straight line, is next determined. This is located by means of scales (one each for ϕ_x, ϕ_y, ϕ_z) which depend on ϕ_k only and the construction of which are discussed in Part II. The scales are actually the abscissae at which the mass distributed centroids (O.F.) lie. For illuminant C and a spectrophotometric curve divided into four segments the ϕ_x scale consists of the four points 447.17, 542.00, 598.03, and 655.46 mμ (see Table IV). The intersections of the linear segments at these points are found and marked. They are the O.F.'s of the segments which by lemma one lie on the segments. The O.F.'s are next connected in pairs (the first and second together, and the third and fourth together) by straight lines. These lines are in reality the linear segments best representing the spectrophotometric curve when divided in halves rather than in quarters. According to lemma two, the O.F.'s of the two quarters and the half are collinear and according to the scale the O.F. abscissae (ϕ_x scale, illuminant C) will lie at 486.21 and 603.74 mμ (see Table IV). The O.F. of the halves are then determined and con-

TABLE IV A. X Scale for C—Illuminant 400–700 Mu.

λ	X_1	X_2	X_3	X_4	X_5
400	414.42				
420					
421	432.09	429.64			
440					
441	450.20		447.17		
460					
461	468.71	456.98			
480				486.21	
481	486.46				
500					
501	515.63	495.35			
520					
521	532.81		542.00		
540					
541	551.69	546.57			
560					565.28
561	571.07				
580					
581	590.59	581.78			
600					
601	610.10		598.03		
620					
621	629.42	617.67			
640				603.74	
641	648.87				
660					
661	668.37	653.53			
680			655.46		
681	688.05				
700		688.05			

TABLE IV B. Y Scale for C—Illuminant 400–700 Mu.

λ	Y_1	Y_2	Y_3	Y_4	Y_5
400	414.52				
420					
421	433.32	431.66			
440					
441	452.05		463.06		
460					
461	471.90	466.05			
480				526.57	
481	491.72				
500					
501	511.55	504.97			
520					
521	531.06		530.95		
540					
541	550.61	541.82			
560					556.95
561	568.27				
580					
581	589.82	577.35			
600					
601	609.64		588.62		
620					
621	629.21	616.40			
640				592.16	
641	648.79				
660					
661	668.35	653.34			
680			655.21		
681	688.03				
700		688.03			

TABLE IV C. Z Scale for C—Illuminant 400–700 Mu.

λ	Z_1	Z_2	Z_3	Z_4	Z_5
400	414.43				
420					
421	432.16	429.76			
440					
441	450.42		449.12		
460					
461	469.29	458.30			
480				454.99	
481	488.54				
500					
501	508.24	492.94			
520					
521	528.47		495.51		
540					
541	547.88	532.17			
560					455.06
561	568.88				
580					
581	589.22	575.39			
600					
601	608.20		579.68		
620					
621	627.66	611.82			
640				579.78	
641	643.83				
648		643.83			
			643.83		

TABLE V A. X Scale for C—Illuminant 380–739 Mu.

λ	X_1	X_2	X_3	X_4	X_5
380	413.39				
420		429.38			
421	432.09				
440			447.03		
441	450.20				
460		456.98			
461	468.71				
480				486.07	
481	486.46				
500		495.35			
501	515.63				
520			542.00		
521	532.81				
540		546.57			
541	551.69				
560					565.34
561	571.07				
580		581.78			
581	590.59				
600			598.03		
601	610.10				
620		617.67			
621	629.42				
640				603.90	
641	648.87				
660		653.53			
661	668.37				
680			656.30		
681	688.05				
700		693.14			
701	711.67				
739					

TABLE V B. Y Scale for C—Illuminant 380–739 Mu.

λ	Y_1	Y_2	Y_3	Y_4	Y_5
380	413.48				
420		431.48			
421	433.32				
440			463.03		
441	452.05				
460		466.05			
461	471.90				
480				526.57	
481	491.72				
500		504.97			
501	511.55				
520			530.95		
521	531.06				
540		541.82			
541	550.61				
560					557.01
561	568.27				
580		577.35			
581	589.82				
600			588.62		
601	609.64				
620		616.40			
621	629.21				
640				592.25	
641	648.79				
660		653.34			
661	668.35				
680			656.03		
681	688.03				
700		693.12			
701	711.67				
739					

TABLE V C. Z Scale for C—Illuminant 380—648 Mu.

	Z_1	Z_2	Z_3	Z_4	Z_5
380	413.41				
420		429.50			
421	432.16				
440			449.00		
441	450.42				
460		458.30			
461	469.29				
480				454.87	
481	488.54				
500		492.94			
501	508.24				
520			495.51		
521	528.47				
540		532.17			454.94
541	547.88				
560					
561	568.88				
580		575.39			
581	589.22				
600			579.68		
601	608.20				
620		611.82			
621	627.66				
640				579.78	
641	643.83				
648		643.83			

nected again by a straight line. The final O.F. abscissa for ϕ_x, illuminant C occurs at 565.28. The ordinate corresponding to this value on the final straight line gives the value of $\bar{R}_x$ which is, from (2), the tristimulus value X corresponding to this curve. Clearly the number of segments into which a curve is split will be determined by the complexity of the reflectance curve. A surprising number of materials may be evaluated by one or two segments only.

More complex curves require 8 foci scales instead of 4 foci scales for their evaluation. The signal glass furnished with many recording spectrophotometers as a calibrating standard gave almost equal values for x when 4 scales and 12 scales were used. In cases where complicated sections of curve arise, it is possible to replace any interval with two half-intervals or four quarter-intervals, representing the corresponding segments of the curve with 2 or 4 straight lines and working in pairs to find an O.F. for that particular original segment. With a little practice, this can be avoided much of the time.

Finding tristimulus values for greys, blacks, whites, or any other curve which deviates but little from a straight line is particularly simple as in this case one single segment suffices and the O.F. will be the intersection of the straight

line with the centroid abscissa 565.28. Table I shows values obtained by three different methods of integration.

In finding the trichromatic coefficients from the data obtained by this method it is first necessary to multiply the three values found from the scales by the ratios

$$\frac{\int_0^\infty \bar{x}E_c d\lambda}{\int_0^\infty \bar{y}E_c d\lambda}=1,\ 1,\ \text{and}\ \frac{\int_0^\infty \bar{z}E_c d\lambda}{\int_0^\infty \bar{y}E_c d\lambda}.$$

These ratios are different for each of the three illuminants.[3] For illuminant A the values are 1.09828, 1, and .35547; for illuminant B .99072, 1, and .85223; and for illuminant C .98041, 1, and 1.18103. The coefficients follow from the normal calculation.

PART II. CONSTRUCTION OF THE SCALES

The scales were constructed from the data and tables given in the *Handbook of Colorimetry*. Inasmuch as the scales are in reality the abscissae of the O.F. they are found by direct numerical integration of Eq. (1) using the appropriate data (Tables XIII, XIV, XV in 1-mμ intervals)[1] and the appropriate limits. The wave-length regions of interest vary somewhat with individual preference; some laboratories are interested in the region 380–770 mμ, others from 400–700 mμ. In order to be sure that no great differences would occur in integrating over either range, scales have been constructed for both ranges for illuminant C. Actually little extra work was involved because of the tabulation method used; the scales will be seen to differ little one from the other. In either case the range was divided into sixteen intervals of approximately 20 mμ each in such a way as to cause the lines of demarcation to fall on easily identified abscissae of the reflectance curve. This particular choice of demarcation was suggested by Dr. D. B. Judd and has materially assisted in the ease of final evaluation. It is not expected that there will be much use for calculations involving sixteen segments but they are given for those few cases where segments must be drawn closely because of some local complexity in the curve. The calculations have been carried out on a Monroe calculator by Mr. C. W. Breuninger and checked independently by Mr. E. M. Man. The following scales have been calculated:

Table II. X, Y, and Z scales for illuminant A wave-length range 400–700 $m\mu$.
Table III. X, Y, and Z scales for illuminant B wave-length range 400–700 mμ.
Table IV. X, Y, and Z scales for illuminant C wave-length range 400–700 mμ.
Table V. X, Y, and Z scales for illuminant C wave-length range 380–739 $m\mu$.

The aid of Mr. J. H. Widmyer and Mr. A. J. Brunner in evaluating the method is gratefully acknowledged.

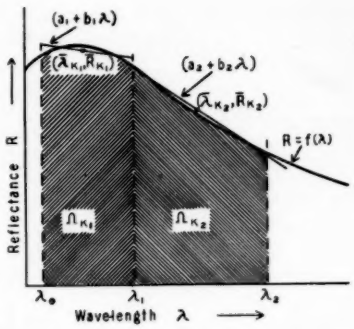


FIG. 3. Schematic reflectance curve. Ω_{k_1} and Ω_{k_2} are actually the areas under the areas under the curves $\phi_k=f(\lambda)_{\lambda_0}^{\lambda_1}$ and $\phi_k=f_1(\lambda)_{\lambda_1}^{\lambda_2}$.

APPENDIX

Lemma 1. *The orthogonality focus of a linear function lies on the curve.* Consider a straight line reflectance curve $R=a+b\lambda$ along which the mass distribution per unit length is given by $\phi_k(\lambda)$. Then, as shown previously, the moment of mass distribution about the λ axis is given by Eq. (2). This is also the expression for the ordinate of the orthogonality focus.

$$\bar{R}_k=\frac{1}{\Omega_k}\int_0^\infty R\phi_k d\lambda.$$

The use of the subscript k denotes either an $\bar{x}$, $\bar{y}$, or $\bar{z}$ luminosity function in determining the product function ϕ_k. Replacing R by its specific

[3] Committee on Colorimetry, J. Opt. Soc. Am. **34**, 643 (1944).

value $a+b\lambda$

$$\bar{R}_k=\frac{1}{\Omega_k}\int_0^\infty (a+b\lambda)\phi_k d\lambda,$$

$$=\frac{1}{\Omega_k}\left[a\int_0^\infty \phi_k d\lambda+b\int_0^\infty \lambda\phi_k d\lambda\right],$$

but by definition and Eq. (1)

$$\Omega_k=\int_0^\infty \phi_k d\lambda$$

and

$$\Omega_k\bar{\lambda}_k=\int_0^\infty \lambda\phi_k d\lambda;$$

hence,

$$\bar{R}_k=\frac{1}{\Omega_k}[a\Omega_k+b\Omega_k\bar{\lambda}_k],$$

$$=a+b\bar{\lambda}_k.$$

Therefore on a linear function the O.F. lies on the curve.

Lemma 2. *If any wave-length interval* (λ_0, λ_2) *of the reflectance curve* $R=f(\lambda)$ *be divided into two parts,* (λ_0, λ_1), (λ_1, λ_2) *the orthogonality foci of the two parts and of the whole are collinear.* (See Fig. 3).

(a). Consider the abscissae $\bar{\lambda}_{k1}$, $\bar{\lambda}_{k2}$ of the orthogonality foci of the two intervals separately. Then from Eq. (1)

$$\bar{\lambda}_{k1}=\int_{\lambda_0}^{\lambda_1} \lambda\phi_k d\lambda \Big/ \Omega_{k1},$$

$$\bar{\lambda}_{k2}=\int_{\lambda_1}^{\lambda_2} \lambda\phi_k d\lambda \Big/ \Omega_{k2}.$$

The abscissa of the orthogonal focus of the total interval $\bar{\lambda}_{kT}$ will be given by

$$\lambda_{kT}=\int_{\lambda_0}^{\lambda_2} \lambda\phi_k d\lambda \Big/ \Omega_{kT},$$

$$=\left[\int_{\lambda_0}^{\lambda_1} \lambda\phi_k d\lambda+\int_{\lambda_1}^{\lambda_2} \lambda\phi_k d\lambda\right]\Big/ \Omega_{kT},$$

$$=\Omega_{k1}\bar{\lambda}_{k1}/\Omega_{kT}+\Omega_{k2}\bar{\lambda}_{k2}/\Omega_{kT}.$$

The abscissa of the orthogonal focus of the whole interval is the linear combination of the O.F. abscissae of the two sub-intervals.

(b) Consider the ordinates $\bar{R}_{k1}$ and $\bar{R}_{k2}$ of the O.F. of the two intervals separately. Then from Eq. (2)

$$\bar{R}_{k_1}=\int_{\lambda_0}^{\lambda_1} R\phi_k d\lambda \Big/ \Omega_{k1},$$

$$\bar{R}_{k2}=\int_{\lambda_1}^{\lambda_2} R\phi_k d\lambda \Big/ \Omega_{k2}.$$

The ordinate of the O.F. of the whole interval will be given by

$$\bar{R}_{kT}=\int_{\lambda_0}^{\lambda_2} R\phi_k d\lambda \Big/ \Omega_{kT},$$

$$=\left[\int_{\lambda_0}^{\lambda_1} R\phi_k d\lambda+\int_{\lambda_1}^{\lambda_2} R\phi_k d\lambda\right]\Big/ \Omega_{kT},$$

$$=\Omega_{k1}\bar{R}_{k1}/\Omega_{kT}+\Omega_{k2}\bar{R}_{k2}/\Omega_{kT}.$$

The ordinate of the O.F. of the whole interval is the linear combination of the ordinates of the O.F. of the two sub-intervals.

Since both ordinates and abscissae of the O.F. of the sub-intervals and the total interval are linearly combined, the three O.F. are themselves collinear.

Photoelectric Properties of Metals in a Finely Divided State

L. J. REIMERT*
Rowland Physical Laboratory, The Johns Hopkins University, Baltimore, Maryland
(Received February 9, 1946)

INTRODUCTION

INVESTIGATIONS, concerning the photoelectric properties of metals, have been restricted, primarily, to clean surfaces prepared either by the conventional method of distillation under the best vacuum conditions obtainable, or by cutting a fresh surface, on the bulk of the metal in a vacuum. These surfaces have in many cases been exposed to various gases, and the effects studied. However, except in such cases as oxidation, the structure of the surface is not radically altered by this procedure.

Pfund[1] has found that when metals such as cadmium, antimony, and zinc are distilled in air under a pressure of several millimeters of mercury, they are deposited, not in the form of the specular surface obtained in a high vacuum, but in the form of a powder which in many cases is intensely black.

The black surfaces were found to have very low reflection and transmission coefficients, the absorption in the spectral range from 0.5 micron to 13 microns being better than 99 percent. Yet because of the extreme fineness of the particles, the surfaces exhibited specular reflection at grazing incidence.

Because of the remarkable optical and physical properties of these surfaces, it was considered of interest to investigate their photoelectric properties. The high optical absorption might increase the photoelectric yield, and, since the specific surface of such surfaces is much larger than in the case of a specular surface, the effects of various adsorbed gases might be enhanced. Furthermore, the extremely small particle size of these films might of itself affect the photoelectric properties.

APPARATUS

The surfaces of cadmium, antimony, and zinc were prepared in a small round bottom flask which had been provided with a suitable distilling unit and a quartz window. For the preparation of the sodium surfaces this cell was altered by the insertion of a glass tube which could be cooled with liquid nitrogen, thus making it possible to deposit the film to be investigated on a cold surface.

The necessary vacuum conditions were obtained by a glass mercury diffusion pump, separated from the cell by a trap which could be cooled with dry ice or liquid nitrogen.

The source of illumination for the examination of the cadmium, antimony, and zinc surfaces was a quartz mercury arc, and for the sodium surfaces a tungsten projection lamp. In the former case the necessary spectral regions were isolated by means of a quartz monochromator, while in the latter case a glass monochromator was used.

Except in the case of sodium, the photoelectric currents were measured by means of a gold-leaf electroscope. The output of the sodium cells was great enough that the current could be determined by a galvanometer amplification system somewhat similar to that of Moll and Burger.[2]

The relative intensity of the lines of the quartz mercury arc, and the energy distribution of the projection lamp, was determined by means of a three-junction, bismuth to bismuth-tin alloy, thermopile kindly supplied by Professor Pfund.

PROCEDURE

The various surfaces were prepared by evaporation from a tungsten wire wound in the shape of a cornucopia. During the process of evacuation the cell was mildly outgassed by sending a high frequency discharge through it, and by heating the tungsten spiral to a temperature somewhat below the melting point of the specimen. After a pressure of about 10^{-6} millimeter of mercury was attained, the current through the spiral was increased until the sample began to evaporate. This procedure produced a specular surface.

The gray and black surfaces were prepared by

* Charles A. Coffin Fellow, 1939–40. Now with the Research Division of The New Jersey Zinc Company (of Pennsylvania), Palmerton, Pennsylvania.

[1] A. H. Pfund, Rev. Sci. Inst. **1**, 397 (1930); J. Opt. Soc. Am. **23**, 375 (1933).

[2] W. J. H. Moll and H. C. Burger, Phil. Mag. **50**, 6211 (1925).

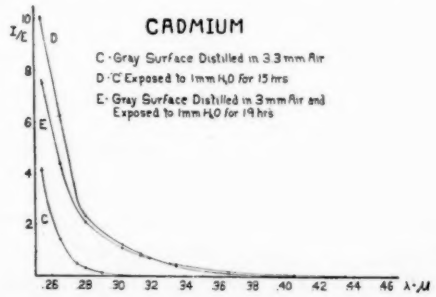
admitting air or argon into the cell to a pressure of from 0.1 to 7 millimeters of mercury and then evaporating until the reflection of the glowing tungsten spiral could no longer be seen on the walls of the cell. The conductivity of these films was so low that, in order to insure good contact with the cathode, they were deposited on top of a specular surface of the same metal. Immediately after these surfaces were deposited, the cell was again evacuated. However, no attempt at outgassing was made.

In order to insure a good contact between the electrodes and the sodium surfaces, the inside of the cell was coated with a film of silver, deposited chemically by the Rochelle salt process, before the sodium was distilled. Furthermore, before the sodium surfaces were deposited, the cell and the remainder of the vacuum system were flamed with a torch so as to promote thorough outgassing.

RESULTS

The results are being presented in the form of curves in which the photoelectric current per unit of incident light intensity, I/E, is plotted as a function of the wave-length of illumination. I/E was determined in arbitrary units; however, the units are the same for all the cadmium, zinc, and antimony curves. The wave-length is plotted in microns.

Many of the results have to do with gray surfaces. These were obtained, in most cases, while attempts were being made to make black surfaces. Even though they do not exhibit the high absorption of "blacks" they are nevertheless interesting because of their extremely small particle size. Furthermore, some of the most interesting results were obtained with them.


CADMIUM
C-Gray Surface Distilled in 3.3 mm Air
D-'C' Exposed to 1mm H2O for 15 hrs
E-Gray Surface Distilled in 3mm Air and Exposed to 1mm H2O for 19 hrs
I/E
λ-μ

. 1. Spectral sensitivity of gray diffusely reflecting cadmium surfaces.

Cadmium

The photoelectric output of a gray surface of cadmium, distilled under a pressure of 3.3 millimeters of air, is shown in curve *C* in Fig. 1. As can be seen by a comparison with curve *A* in Fig. 2, the yield is much less than in the case of a specular surface distilled in a vacuum. Although the thresholds were not determined definitely, due to gaps in the mercury spectrum, it appears from the shape of the curves that the threshold for the gray surface occurs at a somewhat shorter wave-length. Bomke[3] obtained a similar, although smaller, shift when he exposed a specular surface to oxygen.

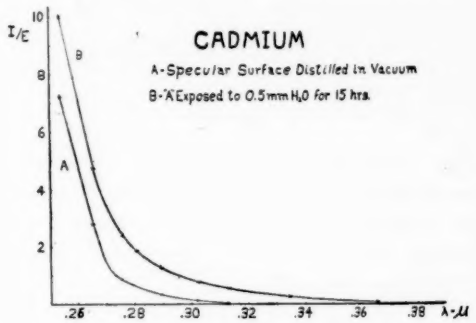


FIG. 2. Spectral sensitivity of specular cadmium surfaces.

When the above gray surface was exposed to one millimeter of water vapor for 15 hours, the improvement shown in curve *D*, in Fig. 1, was obtained. The yield was greatly increased and the threshold was shifted to about 4078A. The shift in threshold was much greater than that of a specular surface similarly treated. The effect of water vapor on a specular surface is shown by curve *B* in Fig. 2. Since the pressure of the water vapor was not critical, a direct comparison can be made between curves *B* and *D*, even though the vapor pressure was not constant.

It was shown that water vapor, and not aging, was responsible for the improvement in photoelectric properties by first aging a gray surface in a high vacuum and then exposing it to water

[3] H. Bomke, Ann. d. Physik **10**, 579 (1931).

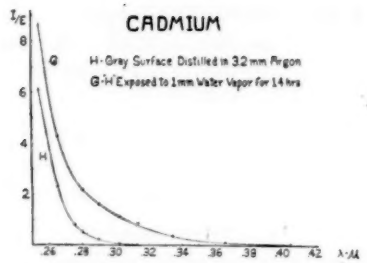


FIG. 3. Spectral sensitivity of gray diffusely reflecting cadmium surfaces.

vapor. After the aging period, no change in the photoelectric properties was observed, while after the exposure to water vapor the usual large increase in yield, and shift in threshold, was observed.

It was also found that the effect thus produced was not permanent. If a surface, sensitized by exposure to water vapor, was kept under a high vacuum for a period of several days, the effect of the sensitization gradually disappeared. However, such a cell could be rejuvenated by a second exposure to water vapor.

Usually thresholds similar to that of curve *D*, in Fig. 1, were obtained after exposing a gray surface to water vapor. However, in several instances, surfaces, which immediately after distillation, gave results similar to *C* in Fig. 1, after exposure to water vapor yielded thresholds similar to that of curve *E* in Fig. 1. Here the threshold was shifted beyond 4358A, and the yield at long wave-lengths is much greater than that usually obtained.

The threshold shift to 4358A or more could not be reproduced at will. Variables whose effects were investigated in an attempt to arrive at conditions controlling the formation of such a surface included: the rate of evaporation, the vacuum conditions, the distance from the tungsten spiral to the surface coated, the temperature of the surface both during the process of distillation and the process of aging, and the position of the surface with respect to the tungsten spiral. Over reasonable limits none of these had a particular effect.

The photoelectric properties of a gray surface distilled in 3.2 millimeters of argon are shown by curve *H* in Fig. 3. The yield is approximately that of a specular surface, and the threshold, though indefinite, probably lies between that of a specular surface and a gray surface deposited in air. This surface, after exposure to one millimeter of water vapor for 14 hours, yielded results similar to those usually obtained from gray air-surfaces similarly treated. The curve thus obtained is *G* in Fig. 3.

A great deal of difficulty was experienced with the preparation of truly black cadmium surfaces. In most cases surfaces, which appeared jet black inside the cell, were found to be only dark gray when examined in the open. It was found that a black surface could best be prepared when the rate of evaporation was extremely rapid. However, under these conditions the pool of molten metal often exploded, so to speak, with the consequence that the metal was spattered, on the wall of the cell, in large lumps. There was some evidence that the presence of water vapor was an aid to the formation of a "black." The pressure of the air during the distillation process was not very important. "Blacks" were formed in air at various pressures from one to four millimeters of mercury. It is interesting to note, however, that no "black" could be prepared in argon.

The photoelectric properties of a black surface distilled in 3.5 millimeters of moist air are shown by curve *E* in Fig. 4. The threshold, as far as

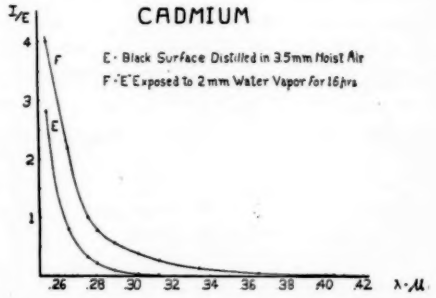


FIG. 4. Spectral sensitivity of black diffusely reflecting cadmium surfaces.

could be ascertained, was the same as that obtained with a gray surface prepared in air. However, the yield was much less. As shown by curve *F* in Fig. 4, after exposure to two millimeters of water vapor for 16 hours, the threshold,

common to most of the gray surfaces prepared in air, was obtained. However, the yield again did not increase to that obtained with gray surfaces similarly treated. The presence of moisture during the process of distillation of black surfaces had no effect on the threshold. That no change was noted is due to the fact that long periods of exposure were found necessary to obtain the maximum effect. Furthermore, the amount of moisture present was relatively small.

Zinc

The output of a gray surface of zinc distilled in 1.5 millimeters of air is shown by curve *B* in Fig. 5. The yield is higher, especially at long wave-lengths, than that of a specular surface, as shown by curve *A*. Also, the threshold has shifted to longer wave-lengths by several hundred angstroms.

A black surface, deposited in 1.2 millimeters of air, gave rise to curve *C* in Fig. 5. To avoid crowding, it is not shown at wave-lengths above 2800A. However, in this region it essentially followed curve *B*, and as far as could be ascertained, had the same threshold.

Water vapor was found to have no effect on the photoelectric properties of the various zinc surfaces. As was the case with cadmium, no black surface of zinc could be prepared in argon. Gray surfaces prepared in argon had essentially the same characteristics as specular surfaces.

Antimony

Contrary to the difficulty experienced in preparing black surfaces of cadmium and zinc, it was found almost impossible to prepare a gray surface of antimony. Dark gray surfaces could

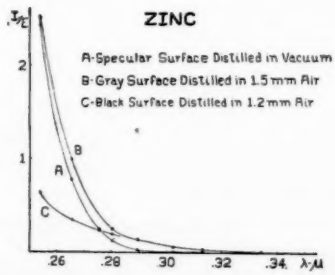


FIG. 5. Spectral sensitivity of zinc surfaces.

be formed in air at pressures above six millimeters of mercury. However, the results obtained with them are of no particular interest.

A black surface distilled in 1.4 millimeters of air gave rise to curve *B* in Fig. 6. Here the yield was, for the most part, greater than that of a specular surface represented by curve *A*. The threshold was also shifted to longer wave-lengths by several hundred angstroms. A black surface prepared in 1.4 millimeters of argon gave rise

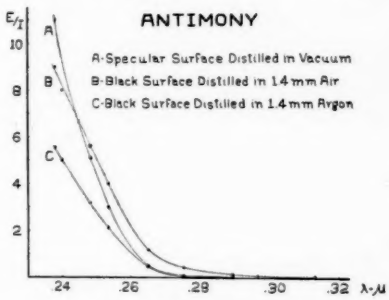


FIG. 6. Spectral sensitivity of antimony surfaces.

to curve *C* in Fig. 6. Except for the yield at short wave-lengths, the output is the same as that of a specular surface. Water vapor was found to have no effect on the antimony surface.

Sodium

The preparation of surfaces by distillation under a fairly high pressure had by far the greatest effect, on photoelectric properties, in the case of sodium. The photoelectric response of a gray surface, distilled, in 1.5 millimeters of air, onto a surface cooled with liquid nitrogen, is shown by curve *B* in Fig. 7. When compared with the output of a surface prepared in vacuum, as shown by curve *A*, the increase in yield is enormous; and not only is the threshold shifted from about 5800A to the neighborhood of one micron, but the emission-maximum is shifted from the ultraviolet to about 5500A. Furthermore, except for a relatively low response in the blue, the sensitivity of the surface is very similar to that of the human eye. As far as could be determined from a run of several days, the effect was permanent.

Attempts to form sodium "blacks" were un-

successful, although Lukens[4] reports that he had evidence of the formation of such surfaces, by distillation, onto a surface cooled with liquid nitrogen, in a fairly high vacuum.

DISCUSSION AND CONCLUSIONS

The interpretation of the preceding results is obviously complicated by the contamination of the surfaces, introduced during the process of their deposition—especially so since the adsorptive power of diffuse surfaces is greater than that of specular surfaces.[5]

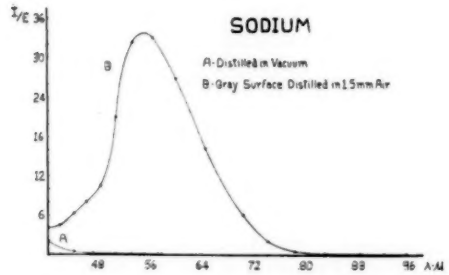


FIG. 7. Spectral sensitivity of sodium surfaces.

The thresholds for specular surfaces of cadmium, zinc, and sodium agree well with those recorded in the literature. A great many of these are tabulated in Hughes and Du Bridge.[6] However, that obtained for antimony is several hundred angstroms lower than the value 3000A to 3100A obtained by Middel.[7]

From the behavior of gray surfaces of cadmium and zinc, it must be concluded that the small particle size of these films has no direct effect on their photoelectric properties. A gray surface of zinc deposited in 1.5 millimeters of air is an improvement over a specular surface, as shown in Fig. 5. However, that this is caused by the effect of adsorbed air is demonstrated by the fact that a gray surface similarly prepared in argon does not show this improvement. The decrease in yield, and small shift in threshold to shorter wave-lengths, of the gray surfaces of cadmium, deposited in air or argon, is caused by the gas adsorbed during the process of deposition. Bomke[3] obtained similar effects by exposing specular surfaces of cadmium to these gases.

The sensitization of cadmium surfaces by water vapor is much greater in the case of diffuse surfaces, both gray and black, than in the case of a specular surface. This is undoubtedly due to the greater adsorptive power of the diffuse surfaces. The nature of the difference in surface structure, which was responsible for the widely separated thresholds of the two types of gray surfaces, after exposure to water vapor, was not established.

The improvement in photoelectric properties of the diffuse surface of sodium over a specular surface, as shown in Fig. 7, is brought about by the influence of air on the surface, and not directly to the structure of the surface itself. However, the sensitization is greater than any previously produced by air. Olpin,[8] by exposing a specular surface to air for a short time, obtained a somewhat greater shift in threshold but the increase in yield was not so great; neither was the large shift of the emission-maximum, to longer wave-lengths, observed. The increase in photoelectric output, obtained in the present investigation, is probably not due to a fundamentally new type of sensitization, but to a different degree of sensitization brought about both by the method of preparation and the structure of the surface.

From the behavior of diffuse black surfaces of cadmium, zinc, and antimony, it must be concluded that neither the small size of the particles comprising these films, nor the high optical absorption, has any direct effect on their photoelectric properties. The best evidence for this lies in the fact that a black surface of antimony, deposited in argon, has the same threshold as a specular surface prepared in a vacuum. The possibility of the argon having a detrimental effect on the film is a remote one. Argon is not adsorbed by these films, to any appreciable extent,[5] and furthermore, it usually does not have much of an effect on photoelectric properties.

The black surfaces of antimony and zinc,

[4] C. Lukens, Dissertation, The Johns Hopkins University, 1938.
[5] L. F. Ehrke and C. M. Slack, J. App. Phys. **11**, 129 (1940).
[6] A. L. Hughes and L. A. Du Bridge, *Photoelectric Phenomena* (McGraw-Hill Book Company, New York), page 76.
[7] V. Middel, Zeits. f. Physik **105**, 358 (1937).
[8] H. R. Olpin, Phys. Rev., **36**. 251 (1930).

deposited in air, show a shift of the threshold to longer wave-lengths; however, this is due to the effect of air on these surfaces during the process of distillation.

The author wishes to express his appreciation to Professor A. H. Pfund for suggesting the investigation and for his continued support, and also to Professor J. A. Bearden for helpful advice.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 36, NUMBER 5 MAY, 1946

Atmospheric Limitations on the Performance of Telescopes *

ARTHUR C. HARDY
Massachusetts Institute of Technology, Cambridge, Massachusetts

CONTRAST REDUCTION BY THE ATMOSPHERE

FROM the standpoint of geometrical optics, a telescope whose magnifying power is M produces an image whose linear dimensions appear to be M times greater than those of the object. It follows that, if the earth had no atmosphere, a perfect telescope would enable the object to be seen at M times its naked-eye range. The purpose of this paper is to show how the known laws of atmospheric optics may be used to predict the optimal performance of a perfect telescope, thus establishing a standard of performance which instruments of this class may approach in practice but can never exceed. The perfect telescope is assumed to provide an image of unimpeachable quality, to possess an exit pupil of adequate size, and to cause no loss of light or degradation of contrast. Its mount is assumed to be stabilized; and the observer experiences no discomfort in using the instrument.

Let a black target whose luminance $B_0=0$ be observed at a distance x along a substantially horizontal line of sight. The differential equations of primary scattering in a homogeneous atmosphere show that the space light between the target and the observer will cause the target to have an apparent luminance (brightness)

$$B_x=B_H(1-\epsilon^{-\beta x}), \tag{1}$$

where B_H is the luminance of the horizon sky and β is the scattering coefficient of the atmosphere. When, in the opposite type of situation, the luminance B_0 of the target is so great that the effect of the space light can be neglected, the apparent luminance of the target decreases exponentially with the target distance because of the atmospheric attenuation. In other words,

$$B_x=B_0\epsilon^{-\beta x}. \tag{2}$$

In the general case, the apparent luminance B_x of a distant target is both increased by the space light and decreased by the attenuation, the net result being

$$B_x=B_H(1-\epsilon^{-\beta x})+B_0\epsilon^{-\beta x}. \tag{3}$$

Let C_0 be the *inherent contrast* of the target seen at a distance $x=0$ against the horizon sky as a background where, by definition,

$$C_0=(B_0-B_H)/B_H. \tag{4}$$

Similarly, let the *apparent contrast* of the same target when observed at a distance x be

$$C_x=(B_x-B_H)/B_H. \tag{5}$$

When Eq. (3) is rewritten in terms of C_0 and C_x, it is found that[1]

$$C_x=C_0\epsilon^{-\beta x}. \tag{6}$$

This means that the contrast of a target seen against the sky along a horizontal path through a homogeneous atmosphere is degraded exponentially with the target distance, regardless of the volume density of the space light.

VISUAL ACUITY ASPECTS

The visibility of a target with the unaided eye depends on many factors, even when a uni-

* Presented at Winter Meeting, Optical Society of America, Hotel Statler, Cleveland, Ohio, March 7–9, 1946.

[1] The simplification resulting from the use of contrast instead of luminance was first called to the author's attention by Dr. S. Q. Duntley.

form target is viewed against a large uniform background. Chief among them are the apparent contrast of the target and the solid angle that it subtends at the eye of the observer. Most of the other factors are irrelevant in connection with the main purpose of this paper because they can be assumed to have substantially the same influence, whether the eye be unaided or be aided by a telescope.

If telescopic vision is compared with unaided vision *at the same target distance*, the benefit that should result from the use of a perfect telescope is adequately represented by the magnifying power of the instrument. That is, the image formed on the retina of the eye is M times larger in linear dimensions when the telescope is used, but is identical with the naked-eye image in all other respects.[2] The highest magnifying power that can profitably be employed under these circumstances is not subject to any simple rule. Preferably, a choice of magnifying powers should be provided, so that the observer can select the power that best meets the requirements of the particular situation.

Binoculars are sometimes used to recognize the presence of a target at a greater range than would be possible with the unaided eye. Unlike the case considered above, the maximum benefit accruing from this use of a telescopic system is readily predictable. Thus, suppose that the target is just visible to the naked eye at a distance r; and that it is likewise just visible through the telescope at a distance R. From Eq. (6) the contrast of the target seen against the horizon sky is less at range R than at range r in the ratio

$$\frac{C_R}{C_r}=\frac{C_0\epsilon^{-\beta R}}{C_0\epsilon^{-\beta r}}=\frac{\epsilon^{-\beta R}}{\epsilon^{-\beta r}}. \tag{7}$$

If the target is to be liminally visible in both cases, the indicated reduction in contrast must be compensated by an increase in the apparent size of the target.

If the target has an area A, the solid angle that it subtends to the naked eye at range r is

$$\omega_r=A/r^2. \tag{8}$$

Correspondingly, the solid angle that the target appears to subtend at range R when using the telescope is

$$\omega_R=M^2A/R^2. \tag{9}$$

The benefit R/r resulting from the use of the telescope evidently depends on the amount of contrast reduction C_R/C_r that the eye will "trade" for an increase in solid-angular subtense ω_R/ω_r.

Many studies of visual acuity are reported in the literature to show how the liminal contrast of a target varies with its angular subtense. From the results of these studies, it is possible to evaluate the basis on which the eye will trade contrast reduction for an increase in the apparent size of the target. It is known, for example, that when the angular size of the target is large (and the contrast correspondingly low), the liminal contrast is almost independent of angular size. This case corresponds to viewing a mountain that is just barely visible to the naked eye through an intervening layer of atmospheric haze. Obviously, the mountain would not be more visible to the naked eye if it were larger; and, for this reason, it could not be seen at a greater distance by means of a telescope, regardless of the magnifying power.

These studies of visual acuity show that the basis on which the eye will trade contrast reduction for an increase in target size becomes increasingly favorable to the use of a telescope as the angular size of the target decreases. In the limit (for small targets of correspondingly higher contrast) the product of target contrast and solid-angular subtense is a constant, a fact which is sometimes known as Ricco's law. It follows therefore that a telescope will perform best when

$$C_r\omega_r=C_R\omega_R, \tag{10}$$

or, on substituting for ω_r and ω_R from Eqs. (8) and (9), when

$$C_R/C_r=R^2/M^2r^2. \tag{11}$$

If this expression is combined with Eq. (7), the following fundamental relationship is obtained:

$$\log M=\log\frac{R}{r}+0.2171(R-r)\beta. \tag{12}$$

[2] It should be noted that the image formed on the retina when the telescope is used is not the same as the image that would be formed if the unaided eye were at $1/M$ the distance from the target. The amount of intervening atmosphere being less at the reduced distance, it is misleading to regard a terrestrial telescope as a device "for bringing objects closer to the observer."

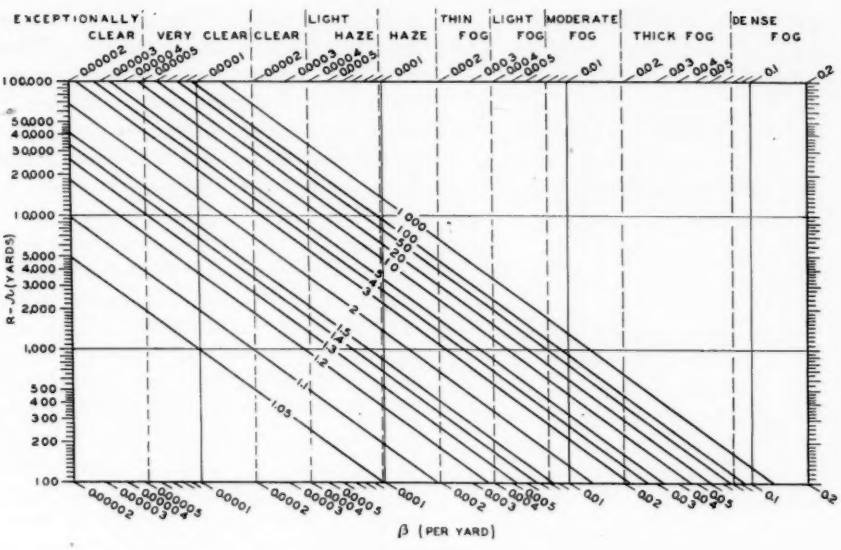


FIG. 1.

If the earth had no atmosphere, β would be zero and Eq. (12) would become

$$M_0 = R/r. \tag{13}$$

In other words, the increase in liminal target range would then be determined solely by the rated magnifying power M_0 of the instrument. Equation (12) may be rewritten in the form

$$\log M = \log M_0 + 0.2171(R-r)\beta. \tag{14}$$

To understand the physical significance of this equation, suppose that a target is just visible to the naked eye at a distance $r = 10{,}000$ yards on a clear day (or night) when the value of β is 0.0003 (per yard). Let it then be required to find the power of a pair of binoculars that would enable the target to be seen equally well at $R = 20{,}000$ yards. In this case, a magnifying power $M_0 = 2$ would provide the eye with an image of the same apparent size as the target seen by the naked eye. However, because of the atmospheric reduction in contrast, an even larger image must be provided; and the second term in Eq. (14) indicates how much larger it must be. On substitution

$$\log M = \log 2 + 0.2171(20{,}000 - 10{,}000)0.0003$$
$$= 0.3010 + 0.6514 = 0.9524,$$

whence

$$M = 8.94.$$

This is the minimum value of the magnifying power in the sense that, unless the target is small, the necessary power would be higher, and possibly very much higher.

GRAPHICAL REPRESENTATION

The relationship expressed by Eq. (14) is more readily grasped if $(R-r)$ is plotted as a function of β for various values of M/M_0, as is done in Fig. 1. Because of the range of the values to be covered, logarithmic scales are desirable for both $(R-r)$ and β. Then, since Eq. (14) can be rewritten in the form

$$\log (R-r) = -\log \beta - \log 0.2171 + \log \log \frac{M}{M_0}, \tag{15}$$

it is evident that the relationship between log $(R-r)$ and log β will be a straight line for constant values of M/M_0. The weather terms at the top of Fig. 1 are those recommended by Hulburt.[3]

[3] E. O. Hulburt, J. Opt. Soc. Am. **31**, 467 (July, 1941).

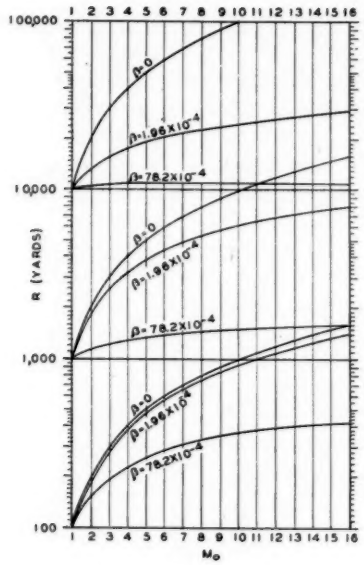


FIG. 2.

As an example in the use of Fig. 1, suppose that the liminal range of a target is 500 yards with the naked eye, and that it is desired to find the power of the glass that will extend the range to 1500 yards. If there were no atmosphere, a 3-power glass would suffice. If the atmosphere is hazy, Fig. 1 shows that an increase of 1000 yards in the liminal range corresponds to $M/M_0 = 2$. In other words, a 6-power glass would be required. This is a minimal value in the sense that, unless the target is small, an even higher power would be required.

Although Fig. 1 represents the optimal performance of optical aids in all kinds of weather, it does not indicate how the liminal range varies with the magnifying power as clearly as does Fig. 2. Here the ordinates are R (on a logarithmic scale) and the abscissae are values of M_0, the rated magnifying power. Three curves are shown in each group, the uppermost curve in each group being the range in a vacuum. The intermediate curve in each group indicates the range on a clear to very clear day (or night) when $\beta = 1.96 \times 10^{-4}$; and the lower curve in each group indicates the range on a day (or night) when there is a light to moderate fog ($\beta = 78.2 \times 10^{-4}$). The three groups of curves illustrate the behavior of optical aids for naked-eye ranges ($M_0 = 1$) of 100, 1000, and 10,000 yards, respectively. Thus, if a signal light is just visible to the naked eye at 10,000 yards on a clear to very clear night, it would be visible at 14,700 yards through a 2-power glass, at 18,200 yards through a 4-power glass, and 22,500 yards through an 8-power glass. It is evident from this figure that, unless the atmosphere is extraordinarily clear or the naked-eye range of the target very short, the law of diminishing returns limits the useful magnifying power to low values.

Figure 2 also indicates the futility of attempting to compare optical aids by means of field tests at liminal ranges. In the case of the example just cited, a 6-power glass and an 8-power glass have liminal ranges of 21,000 yards and 22,500 yards, respectively; and it is doubtful that this difference could be measured by any psychophysical technique that could be employed in field tests. Even if it could, a slight change in the value of β during the test would have a more pronounced effect on the liminal range than either the magnifying power or, in most cases, the optical quality of the instrument under test. It follows that the optical quality of instruments of this class should always be investigated at relatively short ranges indoors where the state of the atmosphere is not a factor and where techniques are available for determining the liminal values (of target distance, size, or contrast) with greater precision.

OBJECTS SEEN AGAINST OTHER BACKGROUNDS

It has been assumed hitherto that the target will be seen against the horizon sky; and it is

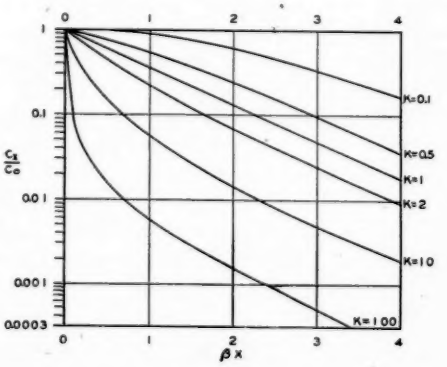


FIG. 3.

not obvious that the same conclusions will apply under other circumstances. Suppose, for example, that the background behind the target is a dark shoreline whose luminance is $1/k$ that of the sky. By following the same line of reasoning that was used previously, it can be shown that the apparent contrast

$$C_x = \frac{C_0 \epsilon^{-\beta x}}{k + (1-k)\epsilon^{-\beta x}}. \qquad (16)$$

This equation takes the form of Eq. (6) when $k=1$, as is to be expected.

Figure 3 shows how C_x/C_0 varies with βx for several values of k. Because C_x/C_0 is plotted on a logarithmic scale, the exponential decrease in contrast that occurs when $k=1$ is represented by a straight line. It will be noted that, when k is greater than unity (background darker than the sky), the rate of decrease of contrast with distance is always greater than the rate for $k=1$, although it becomes equal to the simple exponential rate for large values of βx. Now, a high rate of change of contrast with target distance reduces the benefit that accrues from the use of an optical aid. It follows therefore that, if the target is observed against a background which is darker than the sky, the ratio M/M_0 must be greater than is indicated in Fig. 1. Indeed, for very dark backgrounds and small values of βx, the performance of an optical aid will be decidedly inferior to its performance against a sky background.

Backgrounds brighter than the sky are not often encountered in nature. Nevertheless, it is of interest to note that the curves for $k=0.5$ and $k=0.1$ in Fig. 3 have slopes which indicate a less rapid rate of decrease in contrast than occurs when $k=1$, especially for small values of βx. This means that, in the case of a bright background, the performance of optical aids will be better than in the case of a sky background. As before, the effect of the brighter background disappears when the value of βx is large.

SUMMARY

This paper calls attention to the fact that a telescope may be used (1) to examine an object that can be seen with the unaided eye or (2) to detect the presence of an object that is too distant to be seen with the unaided eye. In the first case, the same amount of atmospheric scattering is interposed between the object and the observer, whether the observer uses telescopic vision or unaided vision. Hence, a perfect telescope provides the retina of the observer with a larger image that is identical with the naked eye image in all other respects.

In the second case, the telescope is called upon to reveal the presence of an object which is beyond the liminal range of the unaided eye. This means that the magnifying power of the telescope must be sufficient to compensate for both the smaller apparent size of the object at the increased distance and the reduction in contrast because of the greater amount of atmospheric scattering. The compensation for the reduction in contrast (by additional magnification) is most effective when the angular subtense of the object is small. It is therefore possible to predict the maximum increase in liminal range that could be secured by a perfect telescope under any assumed conditions of atmospheric scattering. It is shown that the increase in liminal range is small unless the atmosphere is very clear; and that low-power telescopes ordinarily give almost as much increase in range as telescopes of higher power.

Summarizing this summary, terrestrial telescopes enable objects to be seen much better but ordinarily not much farther.

On the Metric of Color Space

ALEXANDER W. WUNDHEILER
Bureau of Ordnance, Navy Department, Washington, D. C.
(Received February 7, 1946)

1. THE PROBLEM

THE available experimental material on color discrimination is not sufficient to define a metric in the color space. It does not even justify the assumption that this metric is riemannian. Nevertheless, some of the foremost physicists[1] have indulged in unbounded speculations concerning this metric, for reasons rather difficult to understand.

It is the purpose of this paper to examine what inferences about the geometry of color space can be made from the known experimental data. Since these data enjoy widely differing degrees of confidence, we shall arrange them in the order of their plausibility and state the geometric properties of color space in the corresponding order, that is: A property No. n is implied by the experimental facts of number n or less.

In numbering our statements, the letters F; H; G; D will denote, respectively: statements of experimental facts; hypotheses; statements about the geometry of color space; properties of the discrimination surface.

2. COLOR SPACE AS A VECTOR SPACE

F1. *Additive color mixing is associative, commutative, invertible (for every pair of colors). The manifold of colors is continuous. Every color can be combined by addition from certain three colors.*

By the very definition of abstract vector space, this implies:

G1. *The color space is a three-dimensional vector space.*

F2. *Colors have a measurable, additive property called luminance; every two colors define a real number, their luminance ratio.*

An additive, scalar vector function is always a linear function $ax+by+cz$ of the components x, y, z of the variable vector. In the centered affine space, commonly used to interpret vectors as oriented segments with origin at the center, its simplest representation is the plane of equation $ax+by+cz=1$, the locus of colors of unit luminance. Colors of luminance l lie in the parallel plane, $ax+by+cz=l$. Thus a special plane direction (not an ordinary direction!) is singled out in the vector space of colors, the plane direction of luminance. Note that there is no natural one-dimensional direction of luminance (the "luminance axis"), contrary to the assumption of some writers.

G2. *The color space contains a special plane direction, the direction of luminance.*

F3. *Colors have a property called chromaticity; colors of equal chromaticity produce by addition a color of the same chromaticity, and of two colors of equal chromaticity one is a multiple of the other.*

This means that colors of equal chromaticity form a one-dimensional vector sub-space of the color space, which is represented by a straight line through the center of the centro-affine space used to represent vectors. Hence:

G3. *The direction of a vector of the color space has the physical meaning of chromaticity.*

Thus far, the only physically significant elements of the vector space of colors prove to be: (a) the origin, (b) the equiluminous planes, (c) the rays through the origin. We shall call them the *singled-out* or *special* elements of the color space.

3. THE PROJECTIVE PROPERTIES OF THE DISCRIMINATION SURFACE

F4. *Discrimination at equal luminance is independent of the luminance.*[2]

[1] H. v. Helmholtz, *Physiol. Optik*, second edition, 444–458; Zeit. f. Psych. Physiol. Sinnesorgane **2**, 1–30; **3**, 1–20, 517 (1891); Akad. der Wiss. Berlin 1071–1083 (1891); Wiss. Abhandlungen **3**, 407–475 (Leipzig, 1895). Erwin Schrödinger, Ann. d. Physik **63**, 481–520 (1920). Ludwik Silberstein, J. Opt. Soc. Am. **28**, 62–85 (1938).

[2] H. Laurens and W. F. Hamilton, Am. J. Physiol. **65**, 547 (1923). W. D. Wright and F. H. G. Pitt, Proc. Phys. Soc. **46**, 459–468 (1934). W. D. Wright, Proc. Phys. Soc. **53**, 96 (1941).
The last paper indicates that a tenfold increase of luminance may decrease the discrimination step fifty percent. G. Haase, Ann. d. Physik **20**, 75–105 (1934), reported a considerable dependence of wave-length discrimination on luminance. His results for $\Delta\lambda$ *vs.* λ differ seriously, however, from the results of Wright and Pitt and the reliability of his method may therefore be questioned.

This means that two just discernible colors of equal luminance remain so when their common luminance is changed without a change in chromaticity. Consider two straight lines joining the origin with two just discernible equiluminous points. Their intersections with any other equiluminous plane will be, according to *F*4, also just distinguishable. Therefore one of these intersections will lie on the discrimination surface of the other. Call the color to which a discrimination surface belongs the *center* of this surface. We see that if the centers of two discriminations surfaces are collinear with the origin, their central equiluminous sections will lie on the same cone with its vertex at the origin. This relationship is called *homothety*. Hence:

D4. *Discriminations surfaces of isochromatic colors have central equiluminous sections homothetic with respect to the origin.*

F5. *Discrimination at equal chromaticity depends on the luminance only.*[3]

This means, geometrically, that if two equiluminous planes cut one isochromatic ray at two just distinguishable colors, they will do so with any other ray. This permits us to speak of *neighboring* (that is, just distinguishable) equiluminous planes.

Every discrimination surface has a special diameter, the one going through the origin; it may be called its *isochromatic* diameter. Thus:

D5. *Isochromatic diameters of discrimination surfaces belonging to equiluminous colors are contained between the same equiluminous planes.*

F6. *Discrimination of luminance is a minimum at equal chromaticity.*[4]

We put *F*6 before *F*7, even though *F*7 is probably the more plausible fact, because *F*6 results in another projective property of the discrimination surface, whilst *F*7 begins the list of its affine properties.

The minimum property stated by *F*6 is meaningful because for every two color-pairs a ratio of "luminance differences" is defined by the two corresponding pairs of equiluminous planes. *F*6 implies that if the direction of the isochromatic diameter is changed slightly, the equiluminous planes through its endpoints come closer together. This entails:

D6. *The tangent planes at the extremities of the isochromatic diameter of a discriminations surface are equiluminous planes.*

4. THE AFFINE PROPERTIES OF THE DISCRIMINATION SURFACE

F7. *The discrimination loci at equal luminance are ellipses of the vector space of colors.*[5]

There is little one could do to continue the discussion of the possible geometry of discrimination, unless *F*7 is construed as evidence for the following statement:

D7. *The discrimination surfaces are ellipsoids of the vector space of colors.*

It will be remembered that the direction of a tangent plane is conjugate to the direction of the diameter going through the point of contact. From *D*6 and *D*7 follows immediately:

D7a. *The isochromatic diameter and the equiluminous plane direction are conjugate for every discrimination ellipsoid.*

It may be of some interest that the last theorem could be obtained from a principle of "economy" for the mechanism of color discrimination:

"The mechanism of color discrimination singles out as few special elements as are necessary to determine the discrimination ellipsoid."

In fact, assuming *D*7, any mechanism of color discrimination would have to assign some conjugate direction to the isochromatic diameter. If this direction were different from that of equiluminance, we would have had, at every point, one more singled-out plane direction, in addition to the direction of luminance. This can be prevented by economically using the already

[3] Hecht, Peskin, and Patt, in J. Gen. Physiol. **22**, 7–19 (1938), showed that this is approximately true for cone vision, i.e., for luminance much greater than 0.01 footlambert. They reported for blue, green, yellow, orange, and red "a maximum variation of about ten per cent from the mean, which is probably of no importance." They showed that the assumption is certainly not correct for rod vision, i.e., for luminance much less than 0.01 footlambert.

[4] No experimental tests of this assumption have been reported. Although it is implicit in most guesses about the way in which sub-liminal differences of luminance and chromaticity are compounded, this assumption has not previously been stated explicitly. It is not easy to devise an experimental procedure adequate for testing this assumption. The ultimate test may have to be indirect, a verification of predictions based on this and better established hypotheses.

[5] L. C. Martin, F. L. Warburton, and W. J. Morgan, Medical Research Council, Special Report Series No. 188 (H. M. Stationery Office, London, 1933). D. L. MacAdam, J. Opt. Soc. Am. **32**, 247–274 (1942). L. Silberstein and D. L. MacAdam, *ibid.* **35**, 32–39 (1945).

existing direction of luminance for one more purpose, that of serving as conjugate to the isochromatic diameter. This results in *D7a* again.

D7a is an experimentally testable proposition since it implies that any equiluminous section of the discrimination ellipsoid should be centered at a point isochromatic with the center of the surface. But no metrical relationship is involved in *D7a*.

The Eq. (*D7b*) expresses the properties *D4* – *D7* of the discrimination surface. To derive it, choose in the vector space of colors a coordinate system (U, V, Y) of origin O, the plane (U, V) or $Y=\text{const.}$ being parallel to the plane direction of luminance.

Consider next a color $M(u, v, y)$ and a local coordinate frame of origin M whose unit vectors are U, V and $OM/y=W$. This last vector is isochromatic and its "luminance span" is 1. We have: $W=uU/y+vV/y+Y$.

If P is a point of the vector space, and the vector MP has the coordinates $\bar{u}'$, $\bar{v}'$, $\bar{w}$ in the triad (U, V, W), it will have the coordinates

$$\bar{u}=\bar{u}'+\bar{w}u/y, \quad \bar{v}=\bar{v}'+\bar{w}v/y, \quad \bar{y}=\bar{w}$$

in the triad (U, V, Y). Indeed, $\bar{u}U+\bar{v}V+\bar{y}Y=\bar{u}'U+\bar{v}'V+\bar{w}W$. The bar refers to the "local" coordinates, that is, with regard to a frame of origin M.

In the frame (U, V, W) the equation of the discrimination surface is, by virtue of *D7*,

$$A\bar{u}^2+2B\bar{u}\bar{v}+C\bar{v}^2+e\bar{w}^2=1,$$

since the plane (U, V) is conjugate to the direction W. Application of *D5* yields at once that *e depends on y only*. In fact, e is the "luminance span" of the isochromatic diameter of the discrimination surface, and if y is constant, so are the equiluminous planes intercepting this diameter.

The central equiluminous section of the discrimination surface has the equation $A\bar{u}^2+2B\bar{u}\bar{v}+C\bar{v}^2=1$. The homothetic section in the plane of unit luminance has, consequently, with respect to the same dyad (U, V), the equation

$$A(\bar{u}y)^2+2B(\bar{u}y)(\bar{v}y)+C(\bar{v}y)^2=1.$$

But this section must be the same for all centers collinear with the origin (by *D4*), that is, for constant u/y, v/y. Therefore, the coefficients of the last equation, $Ay^2=a$, $By^2=b$, $Cy^2=c$ depend on u/y and v/y only.

Using all this, and introducing the local coordinates $\bar{u}$, $\bar{v}$, $\bar{y}$ in the equation of the discrimination ellipsoid, we obtain

$$a\left(\frac{\bar{u}y-u\bar{y}}{y^2}\right)^2+2b\left(\frac{\bar{u}y-u\bar{y}}{y^2}\right)\left(\frac{\bar{v}y-v\bar{y}}{y^2}\right)+c\left(\frac{\bar{v}y-v\bar{y}}{y^2}\right)^2+e\bar{y}^2=1. \quad (D7b)$$

This equation, in the coordinate frame (U, V, Y), is quadratic in the local coordinates $\bar{u}$, $\bar{v}$, $\bar{y}$ of the point of the discrimination surface.

5. THE DISCRIMINATION METRIC

According to its definition, the riemannian discrimination metric is obtained by replacing in the equation of the discrimination surface the "local" coordinates (here $\bar{u}, \bar{v}, \bar{y}$) by the differentials du, dv, dy of the coordinates "in the large," u, v, y:

$$ds^2=a[d(u/y)]^2+2bd(u/y)d(v/y)+c[d(v/y)]^2+edy^2. \quad (M1)$$

Here a, b, c depend on u/y and v/y only, and e depends on y only. Since the form (*M1*) of the ds^2 does not particularly encourage speculation, the natural impulse to introduce the new coordinates,

$$\xi=u/y, \quad \eta=v/y, \quad (1)$$

and obtain the metric

$$ds^2=a(\xi, \eta)d\xi^2+2b(\xi, \eta)d\xi d\eta+c(\xi, \eta)d\eta^2+e(y)dy^2, \quad (M2)$$

is understandable, but must be, nevertheless, carefully controlled: *the coordinates ξ, η may be unacceptable!* This will be the case when the discrimination surface does *not* have the equation

$$a\Delta\xi^2+2b\Delta\xi\Delta\eta+c\Delta\eta^2+e\Delta y^2=1 \quad (2)$$

because Eq. (*M2*) would have no physical meaning. In fact, it could not be derived from the equation of the discrimination surface, as the definition of the discrimination metric requires.

And, indeed, this is the case for the coordinates ξ, η. We have:

$$\Delta u=\xi'y'-\xi y, \quad \Delta v=\eta'y'-\eta y,$$

where $\Delta u, \Delta v$ stand for the former $\bar{u}, \bar{v}$, and the primes indicate coordinates of points on the discrimination surface. We have further:

$$y\Delta u - u\Delta y = \xi' y' y - \xi y^2 - (\xi y y' - \xi y^2) = \Delta\xi y y', \text{ etc.}$$

Substituting this into Eq. (*D7b*) we obtain

$$a\Delta\xi^2 + 2b\Delta\xi\Delta\eta + c\Delta\eta^2 + (e\Delta y^2 - 1)(y/y')^2 = 0. \quad (3)$$

This equation coincides with Eq. (2) if, and only if, y/y' does not measurably differ from 1 in the domain under consideration. In this, and only this case, the coordinates ξ, η are acceptable. Let us remark that the assumption that $y/y' - 1$ is not measurable does not imply that $\Delta\xi, \Delta\eta$ are not either: it is possible to measure an inch along the equator, and still the probable error in any measurement of the latter is greater than an inch. Thus we arrive at the last consequence of the seven "experimental" facts:

In domains where the luminance ratio of two just distinguishable colors is not measurably different from unity, the discrimination metric has the form (M2).

This assumption is satisfied for large luminances. But no further simplification is possible on the ground of present experimental evidence. Recent attempts[6] to do this have been shown to be unsatisfactory.[7]

The footnotes of the present paper have been kindly supplied by Dr. David L. MacAdam to whom the author is also thoroughly indebted for all information concerning the experimental situation.

[6] Parry Moon and D. E. Spencer, J. Opt. Soc. Am. **33**, 260–269 (1943). G. A. Fry, J. Opt. Soc. Am. **35**, 114–135 (1945).

[7] D. L. MacAdam, J. Opt. Soc. Am. **33**, 675–679 (1943); J. Frank. Inst. **238**, 195–210 (1944); J. Opt. Soc. Am. **35**, 616–617 (1945). A. W. Wundheiler, J. Opt. Soc. Am. **35**, 767–771 (1945).

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 36, NUMBER 5 MAY, 1946

Screen-Line Tests of Paraboloidal Reflectors

D. H. HAMSHER
Signal Corps Engineering Laboratories, Bradley Beach, New Jersey

(Received March 11, 1946)

A description of the screen-line test for paraboloidal reflectors is given and the primary considerations involved in its use are presented. The test is useful for reflectors having a face diameter at least as large as four times the focal length. A formula for use in determining small errors in surface regularity is included.

NUMEROUS tests of the optical characteristics of the elements of optical systems have been devised which vary in complexity and in the errors which they measure. The criteria of such a test are: It shall provide information concerning the characteristics which are important from the viewpoint of the intended use and preferably shall be uninfluenced by other characteristics; it should provide easily interpretable results, it being desirable that the results be quantitative and indicative of the location of any fault in the optical element; and, finally, it should be simple to apply and require a minimum of adjustment.

The screen-line test applied to paraboloidal reflectors having a relative aperture $f = 0.5$ used for projection of narrow beams fulfills the above requirements. The test can also be applied to reflectors of smaller numerical aperture by computing the theoretical results and comparing with the experimental results. It is particularly useful when the aperture is small and other methods fail because of the small dimensions involved. This article has been prepared in view of the increasing application of this test and to clarify misunderstandings concerning its proper form and validity which have come to the attention of the author.

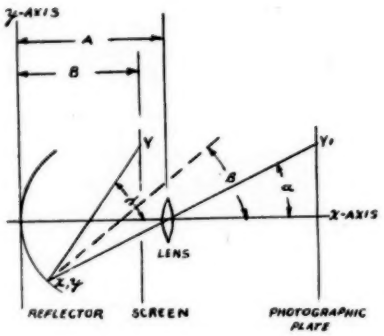


FIG. 1.

I. DESCRIPTION OF THE TEST

The screen, an illuminated surface containing ruled lines, is photographed by reflection from the reflector being tested. The image forming element (later shown to be a pinhole for maximum usefulness) is located on the optic axis of the reflector and in the plane of the screen. The screen is placed at a distance of approximately five times the focal length from the center of the reflector for reasons given below, and perpendicular to the optic axis. The photograph can then be examined for the proper shape of the images of the lines and for the continuity of the lines.

II. SHAPE OF IMAGE LINES

The most general case to be considered is for the condition that the image forming element, temporarily considered to be a lens, is not restricted to the plane of the screen. Figure 1 shows a diagram of the arrangement to which coordinate axes have been added, an arbitrary image point Y has been chosen, and a central image ray has been indicated as being reflected at the point x,y. The dotted line from the point x,y is the normal to the reflecting surface at that point. From considerations of symmetry it is easily seen that the image point Y' must lie in a plane containing the optic axis and the point Y. The ideal arrangement would be one in which the screen and lens were located such that the image of a straight line would be a straight line. Figure 2 represents a projection on a plane perpendicular to the optic axis of an arbitrary line LL' containing the point Y, the

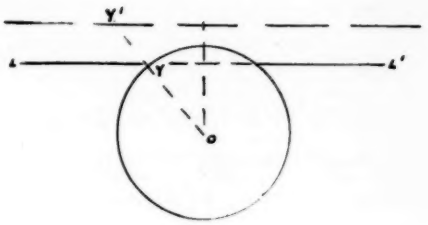


FIG. 2.

center of the reflector O, and the image point Y'. It is obvious that for the locus of Y' to be a straight line the distance OY' must be proportional to the distance OY. This means that the Y' coordinate of Fig. 1 must be proportional to the coordinate Y, or that $\tan\alpha$ is proportional to Y. The mathematical relation between Y and $\tan\alpha$ is derived as follows:

The equation of the parabola is

$$y^2=4fx, \tag{1}$$

where f is the focal length of the parabola. Also

$$\tan\alpha=-y/(A-x), \tag{2}$$

$$\tan\beta=-y/2f, \tag{3}$$

$$\tan\gamma=(Y-y)/(B-x), \tag{4}$$

$$\gamma=2\beta-\alpha. \tag{5}$$

From (4) and (5) we have

$$\tan\alpha=\frac{2\tan\beta-\tan\alpha+\tan^2\beta\tan\alpha}{1-\tan^2\beta+2\tan\beta\tan\alpha}=\frac{Y-y}{B-x}. \tag{6}$$

From (1) and (2)

$$y^2\tan\alpha-4fy-4fA\tan\alpha=0$$

or

$$y=\frac{4f\pm[(4f)^2+16fA\tan^2\alpha]^{\frac{1}{2}}}{2\tan\alpha}.$$

If $\tan\alpha$ is positive y must be negative and therefore only the negative sign is suitable for the problem. Expanding the square root in a series by the binomial expansion theorem, we obtain for the first three terms

$$y=-A\tan\alpha+\frac{A^2}{4f}\tan^3\alpha-\frac{A^3}{8f^2}\tan^5\alpha+\cdots. \tag{7}$$

Substituting (1), (3), and (7) in (6), we obtain

$$Y=-A\left[1+\frac{B(f-a)}{fA}\right]\tan\alpha+\left[\frac{A^2(2f-A)}{4f^2}+\frac{BA(f-A)(4f-A)}{4f^3}\right]\tan^3\alpha$$
$$-\left[\frac{A^3(8f^2-7fA+A^2)}{16f^4}+\frac{BA^2(f-A)(20f^2-10fA+A^2)}{16f^5}\right]\tan^5\alpha, \quad (8)$$

neglecting higher power terms of tan α than the fifth.

The condition that Y is proportional to tan α requires that the coefficients of $\tan^3\alpha$ and $\tan^5\alpha$ both be zero; that is,

$$B=Af(A-2f)/(A-4f)(A-f)$$

and

$$B=\frac{Af(8f^2-7Af+A^2)}{(A-f)(20f^2-10Af+A^2)}.$$

There are no finite values of A and B, except zero, which will satisfy these equations. Therefore, there is no position of the screen and lens such that a straight-line image will result.

Although it is not possible to obtain a straight-line image of a straight line, it is desirable to have the image line as straight as possible. Practical considerations make it desirable also to have the screen and lens in the same plane so that only a small amount of the screen has to be cut away in order to let the light pass through to the camera. Accordingly the problem has been investigated to determine whether or not reasonably straight lines can be obtained if $A=B$. The coefficients of $\tan^3\alpha$ and $\tan^5\alpha$ have been computed as a function of the distance A and are shown in Figs. 3 and 4. The conclusion arrived at above, that the coefficients obtained for $\tan^3\alpha$ and $\tan^5\alpha$ both will not become zero for the same screen distance, is verified in these diagrams.

It is to be noted that, for a given point x,y on the reflector, tan α will be large for small values of A, and the series (7) will not converge as rapidly or will be divergent. Further, it is impractical to employ a screen distance approximately equal to the focal length for large relative aperture reflectors. Consideration of the coefficients plotted in Figs. 3 and 4 shows that at 4.7 times the focal length the coefficient of $\tan^3$

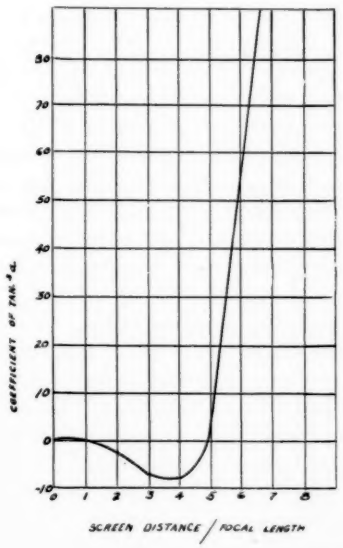


FIG. 3.

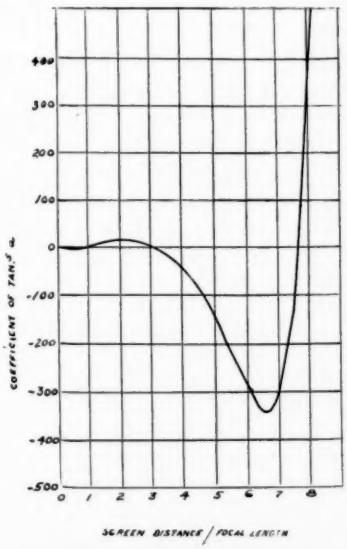


FIG. 4.

α vanishes, and at five times the focal length the second term in Eq. (8) cancels the third term for a value of tan α corresponding to a point on the reflector whose distance from the axis equals the focal length of the reflector. However, at other points on the reflector corresponding to other values the two terms will not cancel. The best screen distance to produce approximately straight-line images with reflectors whose face diameters are equal to or less than twice the focal length is in the neighborhood of 4.5 of five times the focal length. Under these conditions the second and third terms of (8) are so small for the possible values of tan α that for most purposes they can be neglected.

III. REFLECTORS OF LARGE RELATIVE APERTURE

Reflectors having a larger face diameter than twice the focal length require further consideration because the series (7) is convergent only for values of $\tan^2 \alpha$ less than f/A. When the face diameter approaches approximately four times the focal length it becomes necessary to make the screen distance larger than five times the focal length thereby decreasing tan α so that (8) will converge more rapidly for values of tan α corresponding to points near the edge of the reflector. Under these conditions there will be a curvature of all line images over the entire face of the reflector.

Reflectors having a face diameter equal to or less than twice the focal length, if of good quality, produce good straight-line images. However, with reflectors of larger face diameters relative to their focal lengths, larger screen distances must be employed. For example, approximately six times the focal length is a desirable screen distance for a reflector whose face diameter is four times the focal length.

IV. DISCUSSION OF IMAGE FORMING ELEMENT

The ray from the point Y to the point Y' shown in Fig. 1 is not the only ray which contributes to the image Y' if a lens is used. There is a bundle of rays being reflected in the neighborhood of point x,y which passes through the lens and is focused at Y'. Therefore, if a lens is employed, one does not obtain a point correspondence between reflection from the point x,y and the image at Y', and an imperfection in the reflector in the neighborhood of x,y will cause a blurring of the image at Y'. If the lens is stopped down, the region in the neighborhood of the point x,y contributing to the image at Y' decreases, and the image obtained at Y' becomes increasingly representative of the condition of the reflector at the point x,y. Improved characteristics for operation of a test are therefore obtained by the use of a pinhole for the image forming element having a diameter such that the geometrical shadow of the pinhole at the photographic plate is equal to the diffraction pattern of the pinhole obtained at the photographic plate. There is thus obtained, within the limits required for many practical purposes, a point-to-point correspondence between the characteristic of the reflecting surface at a variable point x,y and the corresponding image point Y', which permits, in some cases, errors in the reflector to be traced to the process of manufacture.

V. APPLICATION OF THE TEST

As stated in standard textbooks on optics, the quantity of light received at a distant point from an optical projection system, such as an incandescent lamp and reflector, is proportional to the area of the optical system which is reflecting light in that direction. The particular constants of the source dimensions and reflector are the primary factors which determine the intensity distribution of the projected beam. When there are errors in the reflection surface of sufficient extent such that for a given point no portion of the source is reflected in the given direction, that point will appear dark and not contribute to the illumination. The screen-line test is being applied to reflectors which are to be employed in projection systems for the purpose of rejecting those having too large an area outside the limits of the necessary accuracy of the surface. For a given case, deviation of the lines from a proper position by more than a certain allowed amount is considered a fault in the reflector, and the area of the reflector without the excessive deviation divided by the total area of the reflector gives the fraction of the total possible illumination which will actually be obtained by a properly focused projector. For approximate, quick application of the test the

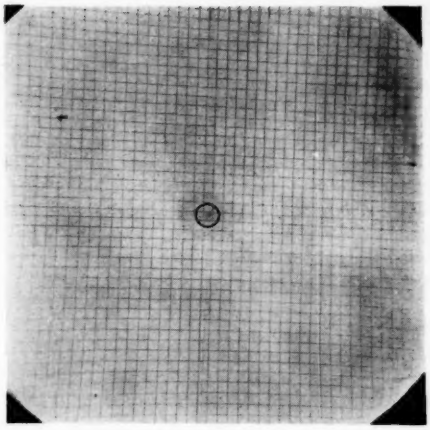


FIG. 5.

counting of squares after the defective areas have been marked out permits a rapid computation of this ratio. Increasing carefulness in determining the deviation of the lines from their theoretical positions gives increasing exactness to the test up to the limit of resolution of the system. If large quantities of reflectors are to be tested, the preparation of templates for rapid determination of defective areas improves the over-all speed in performing the test. In addition to the position of the lines the quality with regard to waviness, fuzziness, contrast, and minute breaks in the line caused by ripples in the surface can be used by experienced testing personnel to detect small errors in the general quality of the surface. Imperfections in the quality of the reflector due to pinholes, minor blemishes, or bubbles which do not affect the total area to an appreciable extent can easily be neglected when examining the photograph.

A quantitative estimate of the angular error in the reflector resulting in a slight waviness of the lines can be obtained as follows. A slight angular error of the surface of the reflector at the point x,y of Fig. 1 will cause a small displacement of the reflection point for the central ray. The geometry is such that small deviations in the angle β are approximately equal to the same deviations in the angle α. The error in the reflector is then given by

$$\epsilon = 56.3\delta \cos^2 \alpha / M,$$

where ϵ is the error in the surface of the reflector in degrees, δ is the deviation from a smooth line, and M is the distance of the photographic plate from the pinhole.

Figure 5 shows a screen-line photograph of a $7\frac{7}{8}$-inch focal length reflector, $f=0.44$. The screen is formed of adjoining ruled sheets in each quadrant and therefore the lines are not continuous across the central lines.

VI. SUMMARY

A discussion of the fundamental characteristics necessary for the employment of a screen-line test to paraboloidal reflectors has been given. It has been indicated how this test provides information concerning the important characteristics of a reflector for use in a projector system, how the results may be interpreted in a reasonably quantitative manner, and how, by proper choice of the image forming element, it may be used to indicate quickly the location of the fault in the optical element. This test has recently been applied on a large scale with considerable success.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 36, NUMBER 5 MAY, 1946

The Trigonometrical Correction of Microscope Objectives

F. D. CRUICKSHANK
Department of Physics, University of Tasmania, Hobart, Tasmania

(Received February 28, 1946)

The general method of differential correction of an optical system developed by the writer is adapted to the particular problem of the trigonometrical correction of a microscope objective. Transfer coefficients are introduced which specify the rate of change of the spherical aberration, the zonal chromatic aberration, the departure from the sine condition, the numerical aperture, and the working distance with the constructional parameters of the system. A numerical example of the correction of a Lister type objective is given.

IN the development of the design of a microscope objective it is usual to correct the system trigonometrically as fully as possible and then to apply the various wave theory tests in terms of path differences. This final trigonometrical adjustment is probably the most laborious part of the work, and also the most empirical, as usually undertaken, for the normal procedure is to change one datum at a time and make repeated ray tracings until the desired correction state is achieved. Thus in the notes of his lectures in 1919 Conrady says: "In arriving at a new design by the method of trigonometrical trials—the only one available for the deep curvatures of microscope objectives—it will be found necessary and also quickest to proceed systematically, changing one datum at a time so as to be able to interpolate in a simple and straightforward way for the desired correction. The temptation will often be strong to superpose a second change, but students may be assured that the process does not pay." The solution of the problem of the transfer coefficients of an optical system given recently[1,2] makes available a new means of carrying out the trigonometrical correction, the new method being free from the empirical features of the old and giving a complete analysis of the potentialities for correction which are present in the system.

THE APPROPRIATE TRANSFER COEFFICIENTS

Suppose that a rough design for an objective of either the Lister or Amici type has been developed in the usual way to a stage at which it is ready for trigonometrical correction. Apart from the chromatic aberration, it is usually sufficient to adjust the design so that the marginal spherical aberration is reduced to zero and the sine condition holds closely. Care must also be exercised during the correction process that the working distance of the objective is not reduced seriously, a fault which may be achieved easily in a high power objective. For these corrections to be undertaken it will be adequate to trace a paraxial and a marginal ray from an axial point determined by the tube length required, the trace being made towards the side of the shorter conjugate. The paraxial ray is traced for convenience with an initial inclination angle $u=\sin U_m$, and then the departure from the sine condition may be measured by

$$\epsilon=\sin U_{mk}'-u_k'. \tag{1}$$

For each constructional parameter, α_i, which may be a curvature, thickness, or refractive index, we can calculate, by means of expressions given previously by the writer,[2] the transfer coefficients

$$\frac{\partial L_{mk}'}{\partial \alpha_i}=C(\alpha_i)_m \csc U_{mk}', \tag{2}$$

$$\frac{\partial l_k'}{\partial \alpha_i}=C(\alpha_i)_{pz}/u_k', \tag{3}$$

$$\frac{\partial}{\partial \alpha_i}(\sin U_{mk}')=\cos U_{mk}'\frac{\partial U_{mk}'}{\partial \alpha_i}, \tag{4}$$

whence

$$\frac{\partial LA_m'}{\partial \alpha_i}=\frac{\partial l_k'}{\partial \alpha_i}-\frac{\partial L_{mk}'}{\partial \alpha_i}, \tag{5}$$

$$\frac{\partial \epsilon}{\partial \alpha_i}=\frac{\partial}{\partial \alpha_i}(\sin U_{mk}')-\frac{\partial u_k'}{\partial \alpha_i}. \tag{6}$$

[1] A. L. McAulay and F. D. Cruickshank, Proc. Phys. Soc. **57**, 302 (1945).
[2] F. D. Cruickshank, Proc. Phys. Soc. **57**, 350 (1945).

TABLE I.

Surface		L_{mk}'	i_k'	sin U_{mk}'	u_k'	LA_m'	ϵ
1.	$\partial/\partial c$	−310.63	−273.27	0.46527	0.35347	37.36	0.11180
	$\partial/\partial d$	0.0155	0.0163	−0.00123	−0.00123	0.0008	0.00000
2.	$\partial/\partial c$	−67.38	−50.70	0.10885	0.09412	16.68	0.01473
	$\partial/\partial d$	0.2779	0.2320	0.00335	0.00341	−0.0459	−0.00006
3.	$\partial/\partial c$	312.59	294.19	−0.68287	−0.66884	−18.40	−0.01403
	$\partial/\partial d$	0.1418	0.1277	0.00242	0.00252	−0.0141	−0.00010
4.	$\partial/\partial c$	−91.84	−89.18	1.51296	1.49435	2.66	0.01861
	$\partial/\partial d$	0.2306	0.2070	0.00394	0.00408	−0.0236	−0.00014
5.	$\partial/\partial c$	−20.56	−14.86	0.35174	0.26936	5.70	0.08238
	$\partial/\partial d$	0.8736	0.8283	−0.00215	−0.00162	−0.0453	−0.00053
6.	$\partial/\partial c$	87.64	80.65	−1.59812	−1.54060	−6.99	−0.05752
	$\partial/\partial d$	0.6076	0.6169	0.00000	0.00000	0.0093	0.00000

Equations (5) and (6) provide the necessary transfer coefficients for the adjustment of the spherical aberration and the sine condition, while Eq. (3) gives the rate of change of the working distance, and Eq. (4) the rate of change of the numerical aperture, with the parameter α_i. If additional traces of zonal rays for *C* and *F* light are made, transfer coefficients specifying the variation of the chromatic aberration with the constructional parameters may be computed. Thus

$$\frac{\partial Lch_z'}{\partial \alpha_i} = \frac{\partial L_{zC}'}{\partial \alpha_i} - \frac{\partial L_{zF}'}{\partial \alpha_i}, \tag{7}$$

the transfer coefficients on the right of Eq. (7) being computed from relations similar to Eq. (2). The calculation of these transfer coefficients provides the designer with a complete analysis of the correction potentialities of the system. This is best demonstrated by the consideration of an actual example.

NUMERICAL EXAMPLE—LISTER TYPE OBJECTIVE

In the course of developing a design for a Lister objective a stage was reached at which the system had the residuals in the first line of Table II. To examine the correction potentialities of the objective the transfer coefficients shown in Table I were calculated. The dimensions of the system were given in millimeters so that the curvature unit employed is the reciprocal millimeter. According to the first line of Table I, for example, a curvature change of 0.001 mm^{-1} at the first surface will cause a shift of the final intersection point of the marginal ray through 0.311 mm, a change of spherical aberration by 0.037 mm, and a change of the sine condition aberration quantity, ϵ, by 0.000112, etc. According to the eighth line of the table a shift of the fourth surface together with all surfaces behind it—thereby altering the central airspace of the system—through a distance of one millimeter will introduce a spherical aberration change of amount 0.024 mm and an ϵ-change of 0.00014. Table I thus gives the exact effect of a small change in curvature at any surface of the system and the effect of a small change in the thickness of any component or airspace. For the sake of brevity the transfer coefficients for refractive index changes have been omitted from the example. A study of this table enables the designer to select which changes may be made for the trigonometrical correction of the objective. Obviously, this may be done in a number of different ways. In this particular system it is easier to obtain the spherical correction than the ϵ-correction. The change which offers the greatest correction for the sine condition, a shift of the fifth surface, cannot be used to advantage in the present case as the direction in which the change must be made to improve the correction results in the thinning down of a biconvex lens which is already close to its minimum thickness for reasonably easy manufacture. One simple change which would remove the spherical aberration and reduce the ϵ-value considerably is a positive *d*-change at surface four. A short calculation shows that if this surface were shifted

TABLE II.

	L_{mk}'	l_k'	sin U_{mk}'	u_k'	LA_m'	ϵ
From original trace	11.339	11.396	0.21820	0.21769	0.057	0.00051
Predicted changes to first order	0.558	0.501	0.00953	0.00987	−0.057	−0.00034
Predicted changes corrected	0.534	0.478	0.00953	0.00987	−0.056	−0.00034
Predicted values new trace	11.873	11.874	0.22774	0.22756	0.001	0.00017
Actual values new trace	11.878	11.875	0.22769	0.22757	−0.003	0.00012

forward through 2.42 mm in the usual way the spherical aberration will be reduced to zero and ϵ will be reduced to one-third of its present value. The predicted results of this change are entered in the second line of Table II. Correcting these expected changes by the use of a second-order correction term,[3] for a displacement of 2.42 mm is scarcely a differential change in a system of these dimensions, we obtain the third line of Table II. The fourth line of Table II combines the original values of the various quantities with the predicted changes and the last line gives the values of these same quantities as shown by a fresh trace of the altered system. The agreement with expectation is of a very high order, and demonstrates the usefulness of the new method of correction.

It is not to be expected that the agreement with prediction will always be as good as that given in the example, particularly if a large number of changes are made simultaneously. In general, however, it is sufficiently good to require only a further small adjustment after the new trace is made. Looking on the matter geometrically, if a curve is drawn representing the relation between the aberration A and some constructional parameter α_i of the system, then the slope of the tangent to the curve at a point is specified by the corresponding differential coefficient $\partial A/\partial\alpha_i$. If the slope of this curve is changing slowly in the vicinity of the point representing the present value of the aberration in the system, the value of the aberration change predicted by the transfer coefficients for a small change in the parameter will be very close to the actual value. If, however, the slope of the curve is changing rapidly in the vicinity of this point, the agreement of the actual aberration change with that predicted by the transfer coefficients will obviously not be so close.

[3] F. D. Cruickshank, Proc. Phys. Soc. **57**, 362 (1945).

Modified Rayleigh Scattering in a Liquid

D. H. RANK
The Pennsylvania State College, State College, Pennsylvania
(Received February 25, 1946)

L. BRILLOUIN[1] by introducing the Debye elastic heat waves into the theory of scattering of light predicted the existence in the scattered light of modified frequencies differing in wave-length from the incident radiation by an amount

$$d\lambda = \pm 2\lambda \frac{v}{c/n} \sin \frac{\varphi}{2},$$

where λ is the wave-length of light, v the velocity of sound in the medium, c the velocity of light, n the index of refraction, and φ the angle of scattering.

E. Gross[2,3] first experimentally demonstrated the existence of the above-mentioned phenomenon. W. Ramm[4] has made a rather complete experimental and theoretical study of this subject. Ramm has largely cleared up many of the conflicting bits of experimental evidence on this phenomenon which were observed by Gross and others. A complete bibliography of the subject up to the time of Ramm's paper is to be found in reference 4. The observation of the Brillouin effect as can be seen by reference to the above equation requires the use of apparatus of extremely high resolving power. Gross and Ramm both used echelon gratings as the high resolving instruments. Ramm[4] made use of both λ4358 of mercury and λ4680 of zinc. Ramm showed that many of the conflicting results in the literature with regard to observations of this phenomenon could be explained by the troublesome hyperfine structure of the mercury lines when examined in conjunction with interference spectroscopes. For this reason much of his work was done with λ4680 of zinc despite the troublesome nature of the zinc vapor lamp.

In 1936 the author[5] reported some experiments on the Brillouin effect in a paper delivered at an American Physical Society meeting. Because of neglect on the part of the author these results were never published. The experimental investigation of this phenomenon is still incomplete in spite of the excellent work of Ramm.[4] It would appear that experimental work on this subject with totally different apparatus than that used by all other investigators is fitting and proper.

EXPERIMENTAL

Light Source

The experimental arrangement used to produce the scattered light was similar to the "light furnace" described by R. W. Wood.[6] The scattering tube was of Pyrex about 115 cm long and 37 mm in diameter. At one end of the tube a plane Pyrex window was sealed on to the tube. A cone shaped extension of the tube bent smoothly until the end portion formed a right angle with the length of the scattering tube formed the basis of a black background. The scattering tube was provided with a water jacket (54 mm in diameter), in order to prevent heating of the liquid. The mercury arcs were constructed of Pyrex glass 37 mm internal diameter. These arcs burned between mercury pool electrodes 50 cm apart. In order to maintain low pressure in the arcs they were operated on a high vacuum system with reflux condensers in the vacuum line. A current of about 10 amperes was passed through each arc. Arcs and scattering tube were surrounded by sheet aluminum reflectors. Proper precautions were taken to reduce parasitic light from the scattering tube to a minimum. Light from the scattering tube was focused on the slit of a grating spectrograph with an achromatic lens. The geometry of the scattering apparatus is such that by far the major portion of the scattered light observed was scattered through an angle of 90° with respect to the incident light.

Spectrograph

The spectrograph consisted of a 21-foot concave grating in an Eagle mounting. The grating

[1] L. Brillouin, Ann. d. physique **17**, 88 (1922).
[2] E. Gross, Naturwiss. **18**, 718 (1930).
[3] E. Gross, Nature **126**, 201, 400, 603 (1930).
[4] W. Ramm, Physik. Zeits. **35**, 756 (1934).
[5] D. H. Rank, Phys. Rev. **49**, 880 (1936).
[6] R. W. Wood, Phil. Mag. **35**, 236 (1918).

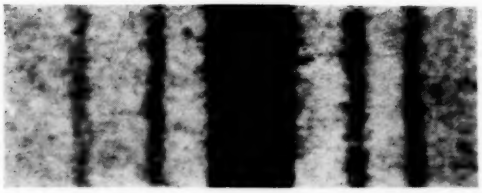


FIG. 1. Seventy-fold enlargement of λ4358 photographed in the 5th order of 21-foot grating.

was one ruled by Gale at the University of Chicago. Ruling was on speculum metal 600 lines per mm and $5\frac{1}{4}$ inches wide. All orders up to and including the 5th are usable and of moderate intensity. The grating end of the spectrograph was in a room the temperature of which was controlled better than .1 degree C. The grating holder, etc. were mounted on a massive carriage (old milling machine bed) and these parts enclosed in a galvanized iron box inside the constant temperature room. This arrangement allowed the temperature of the grating to be maintained constant to much better than .1 deg. C. The limiting factor on exposure time with this spectrograph when high resolution was demanded was always the changes in the barometric pressure and not temperature control. Plates taken on the CuO red bands with a 10-hr. exposure in the 3rd order gave resolving power of 190,000. A third-order plate of the 0→2 band of N_2 first positive system (λ7750) taken with a 20-hr. exposure also showed approximately 200,000 resolving power. These examples are cited to show that the temperature control of the spectrograph was adequate to maintain the resolving power of the spectrograph over an exposure time of considerable duration. A 70-fold enlargement of λ4358 of mercury obtained in the 5th order of the grating is shown in Fig. 1. This photograph was made on a fine-grained emulsion using light from the arcs described above. The slit width was .015 mm, exposure 1 hr. From the width of the lines the resolving power must be above 300,000. For comparison purposes the hyperfine pattern of λ4358 partially observed and partially calculated by Schüler[7] and Jones is given in Fig. 2. The components labeled *A*, *B*, *C*, *D*, *E*, and *F* can be seen to appear on the grating plate Fig. 1, with approximately the relative intensity predicted by Schüler and Jones.

Plates of the Rayleigh scattering from benzene were obtained in the 4th order spectrum of the grating discussed above. The slit width used was .025 mm. Exposure times varied from 9 to 12 hours. If the barometric pressure varied by more than 2 mm during the time of the exposure, the plates were discarded. Blank plates were taken of the direct light of the arc scattered from white paper. These blanks required about 1-hour exposure. The resolving power obtained was in excess of 200,000.

The plates were microphotometered using a Leeds and Northrup microphotometer. Reproductions of the microphotometer traces are shown in Fig. 3. *A* and *C* in Fig. 3 are microphotometer traces of the λ3650 and λ4358 lines of mercury scattered from benzene. *B* and *D* are the microphotometer traces of these same lines scattered from white paper. The presence of the symmetrically placed partially resolved Brillouin components can clearly be seen on the microphotometer traces.

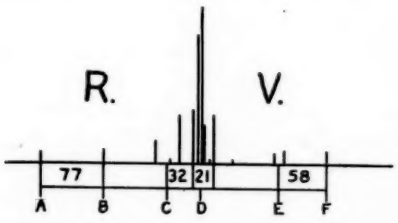


FIG. 2. Hyperfine structure of λ4358 of Hg according to Schüler and Jones. The numbers in the boxes are separations in milliangstroms. *A*, *B*, *C*, *D*, *E*, and *F* refer to components observed in 5th order grating plate Fig. 2.

The microphotometer curves, Fig. 3, are very similar to those reproduced in Ramm's[4] paper, e.g., curves obtained with his small echelon in the single order position with 180° scattering. The beautiful double-order plates obtained by Ramm, with his large echelon show greatly enhanced resolution of the Brillouin components because of the operation of the intensity distribution in the two orders. Measurements of the separation of the Brillouin components as made from the

[7] H. Schüler and E. G. Jones, Zeits. f. Physik **74**, 639 (1932).

author's microphotometer traces yielded the following $d\lambda$ for benzene.

$$\lambda 3650 \quad d\lambda = .034A \pm .003,$$
$$\lambda 4358 \quad d\lambda = .039A \pm .003.$$

The values calculated from Eq. (1) were .033A and .039A, respectively, for λ3650 and λ4358.

The microphotometer traces clearly show the existence of a general continuum in the neighborhood of the Rayleigh line scattered from benzene. The existence of this general continuum is well known.[5] In the case of scattering from benzene a strong general continuum extends many angstroms from the exciting line so that this continuum is readily observable even with low dispersion instruments.

The measurements made by Ramm[4] and those of the author show that the unmodified line is somewhat stronger than the Brillouin components. The experimental fact that the above situation is the true state of affairs rests on very insecure grounds. The experiments conducted by the author and those of Ramm[4] both took reasonable care to avoid dust in the liquid and eliminate parasitic light. From the author's experiences with the Raman effect it can be said

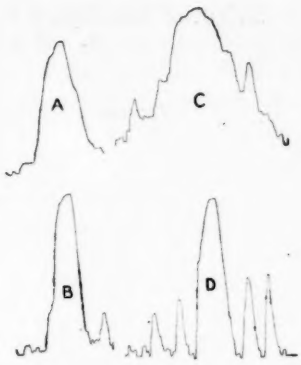


FIG. 3. *A* and *C* microphotometer traces of λ3650 and λ4358 scattered from benzene. *B* and *D* microphotometer traces of λ3650 and λ4358 scattered from white paper.

that the experimental conditions used in the above work with regard to elimination of dust and parasitic light were such that beautiful low dispersion Raman spectra could have been obtained. The intensity of the parasitic light, however, was not reduced to the minimum possible and hence the intensity of the unmodified line shown on the plates is an upper limit if it results from true scattering from the liquid.

The author wishes to thank Mr. S. W. Harding for making the microphotometer traces.

[5] W. Gerlach, Ann. d. Physik **1**, 30 (1929).

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 36, NUMBER 5 MAY, 1946

Effect of Modifying the Spherical Aberration on the Resolution of a Lens Erecting Telescope

D. H. RANK*
Instrument Division, Frankford Arsenal, Philadelphia, Pennsylvania
(Received March 15, 1946)

THE introduction of exit pupils much larger than the normal daylight pupil of the eye in low power telescopic systems gives rise to some formidable optical design problems. In these low power systems the demand for a large exit pupil is also usually coupled with demand for a large field of view. The design problem which must necessarily be a compromise solution becomes truly formidable. The general use of anastigmatic optics in telescopic systems of the type mentioned above has produced telescopes of this type much superior to any available in the early days of the war.

This work was undertaken with the idea of improving the performance of one of the earlier telescopes designed at Frankford Arsenal before it was known what the best balance of aberrations should be to give the optimum performance throughout the field of view. The telescope in question was a three-power instrument with lens erection which had a 7-mm exit pupil. The design was afflicted with approximately between five and six Rayleigh limits of spherical aberration for the completed instrument. This spherical aberration was not of the zonal variety but practically pure undercorrection.

With the power and exit pupil diameter given above it can be seen that the axial resolving power of the complete system eye plus telescope was limited only by the resolving power of the eye since the pupil of the eye acted as a stop in the system. Under normal or bright daylight conditions the pupil of the eye will have a diameter of 3 or 4 mm. Over a pupil diameter of this magnitude the telescopic system will be well within the tolerance for spherical aberration and the imagery will be much better than the eye can make use of at the magnification employed. However, if the pupil of the eye be moved around rather rapidly over the area of the exit pupil of the instrument an annoying "flicker" occurs. This "flicker" is most noticeable when the illumination is poor (a dreary rainy day), and the pupil of the eye has expanded somewhat over its size in bright daylight. Knowing the spherical aberration characteristics of the telescopic system makes it immediately obvious that the "flicker" results from a combination of the aberrations in the exit pupil and the attempt of the eye to accommodate for the slightly different focus in different parts of the pupil.

It was known that the spherical aberration present in the telescope was produced solely by the second erecting lens. This second erector was of the type first designed by König[1] and consisted of a single element airspaced from a cemented doublet. See Fig. 1. The idea occurred to the author that figuring this lens to substantially remove the spherical aberration would remove the "flicker" in the telescope.

The aberrations in an optical system as is well known are not independent of each other. Since the off-axis aberrations in general have much larger values than the axial aberrations, it was thought that correction of the axial spherical aberration by figuring would remove the axial "flicker" of the telescope without materially changing the off-axis defects for better or for worse. The results given in the experimental part of this paper will show that within the limit of experimental error the above statement is true.

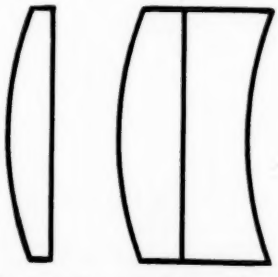


FIG. 1. Schematic diagram of lens which was figured.

* Now at The Pennsylvania State College, State College, State College, Pennsylvania.

[1] A. König, U. S. Patent, No. 2,158,507.

No trigonometrical investigation of the effects of figuring this lens with respect to the off-axis aberration was carried out. Recently a large lens of high relative aperture has been designed by the author in conjunction with Mr. George Arnold. The spherical aberration of this lens is of the usual zonal type commonly associated with high relative aperture. A trigonometrical investigation of the effect of figuring the last surface of this lens to remove the zonal spherical aberration has been made. The results show that the primary sagittal coma, as indicated by the optical sine theorem, will be lessened by the figuring process.

EXPERIMENTAL

Notes on axial testing with the lens interferometer

Twyman[2] has shown how it is possible to obtain quantitative values of axial and off-axis aberrations from the interference patterns with the aid of his "deflector" system.

When a perfect lens is placed on the interferometer and brought into adjustment so that the focus is exactly at the center of curvature of the convex mirror the field of the instrument will be a solid color over the entire lens. If the convex mirror is tilted slightly a number of perfectly straight interference fringes cross the field. If a slight amount of spherical aberration is present in the lens, the tilted mirror will reveal that the fringes are not straight lines but curved near the edge of the lens. If the aberration is not too bad, it may still be possible to orient the convex mirror so that the whole field appears to be a solid color. When the above condition is fulfilled, the lens has an optical path error of less than $\frac{3}{8}\lambda$. Such a lens "green" lens in contradistinction to a perfect lens or "black" lens closely approaches the Rayleigh $\lambda/4$ criterion for image perfection. In the case of primary spherical aberration Conrady[3] has shown that O.P.D. is a minimum when the image produced by such a lens is examined at a point midway between the geometrical focus for the paraxial and the rim zone. When such a lens is placed in the lens interferometer and adjusted with respect to the convex mirror so that the largest possible solid green circular area in the center is obtained, we have achieved the experimental position for the minimization of the O.P.D. for a lens of diameter equal to the diameter of the solid green area. In the presence of considerable spherical aberration a number of rings appear surrounding the solid green area of maximum size. Experiments made with simple crown lenses whose spherical aberration was purely primary have shown that in this simple case when the lens is set in the maximum solid green position the number of rings appearing corresponds to the number of Rayleigh limits of spherical aberration present, the tolerance being of course determined for the rim ray. It is seen that the longitudinal value of the spherical aberration is thus easily obtained when the aberration is primary.

Figuring

The optical figuring procedure consists in removing some glass from the center of the lens relative to the edge. The Twyman lens interferometer furnishes a very satisfactory device for testing the optical path contour of the lens. From inspection of Fig. 1 it can be seen that the figuring process could be applied to anyone or all of the four air glass interfaces. For axial image points the resolution would be the same regardless of which surfaces or surface had been figured since perfection will depend solely upon fulfilling the condition that the optical paths through all zones of the lens are equal (i.e., a plane wave transformed by means of the lens into a truly spherical wave front). Figuring to remove the spherical aberration was accomplished by hand polishing. The polishing tool was made of medium hard pitch slightly smaller in diameter (tool diam. $1\frac{1}{8}$ inches) than the lens being figured. A central portion of the tool approximately $\frac{1}{3}$ the tool diameter was relieved so that no polishing occurred from this region of the tool. The polishing stroke used was the same as that normally used in figuring a parabolic mirror. Polishing during figuring was performed using a very considerable pressure so as to rapidly "wipe" off the glass in the desired region without working the surface too generally. The purpose of the large relief in the center of the tool was to

[2] F. Twyman, Phil. Mag. **42**, 787 (1921).
[3] A. E. Conrady, *Dictionary of Applied Physics*, edited by Sir Richard Glazebrook (The Macmillan Company, Ltd., London, 1922–1923), Vol. IV, pp. 218.

prevent polishing a "hole" in the center of the lens. It was possible to correct the spherical aberration of these lenses using any one of the air glass surfaces as the object of the figuring process. The surface which seemed to yield results most easily was the strong negative curve on the flint of the cemented doublet shown in Fig. 1. The correction achieved by the author in figuring these lenses was never complete. The best that was done was to reduce the spherical aberration from 6 rings to 1 ring as measured by the interferometer. It was thought that the difficulty encountered with the last ring was probably caused by a turning down of the edge of the lens while figuring. However since the author spent only two or three days in devising strokes cuts and figuring about a dozen lenses the technique used could probably be considerably improved.

Figuring of lenses of the type described above is very easy and requires no great skill. As an example one of the author's assistants who had never polished a piece of glass previously was able to satisfactorily correct one of these lenses after less than a day of observation and instruction. This man however was familiar with the use of the interferometer and the interpretation of the interference patterns. The author's experience in this and in other instances is in complete agreement with the statement of Twyman[4] with regard to the simplicity of executing certain figuring operations when the interferometer is used as the control instrument.

Resolution of Telescopes

The method used to measure the resolution of these telescopes was as follows. The target was the familiar Foucault test object. Distance between target and telescope could be continuously varied and accurately measured so that the resolution was directly proportional to target distance at which resolution occurred. It was found that measurements of target distance could be reproduced within the focal range of the telescope objective as measured in the object space. Measurements were made using an auxiliary telescope in back of the eyepiece of the instrument. The magnification of the auxiliary telescope was sufficient so that the physical resolving power of the telescope under test and not the eye of the observer was the limiting factor for resolution. The author has shown by trigonometrical investigation of a similar telescope that the target distances used were sufficiently large so that the finite target distance does not change the spherical aberration characteristics of the telescope by more than a small fraction of the Rayleigh tolerance.

When off-axis measurements of resolution of a telescope are made, it is necessary to set up criteria as to what is meant by resolution due to the inevitable presence of astigmatism and curvature of field. The criterion used for astigmatism was that the resolution distance was taken as the distance at which either primary or secondary images ceased to be resolved, whichever occurred first. Curvature of field was ignored since the eyepiece of the instrument under test was focused to give the best images possible. In addition the curves shown in Fig. 2 are very slightly in error because of the fact that field of view was measured by the rotation of the telescope about its geometrical center rather than about the entrance pupil of the objective. The criteria made use of in off-axis testing however have no effect on the results given in this paper since astigmatism and curvature of field will be sensibly constant with the two types of erecting lenses used. It might be pertinent to point out that the use of the above-mentioned criteria for off-axis resolution is not strictly correct when comparing anastigmatic designs with non-anastigmatic designs. It is well known that the eye can tolerate a certain amount of curvature of field and astigmatism.

For comparison purposes it seems desirable to compare the performance of a telescope with that of a perfect instrument. Conrady[5] has discussed the resolving power of telescope objectives. He shows that the expression formulated by Abbe for the resolving power of the microscope applies to telescopes as well and yields for stars a result in agreement with Dawes' observations on telescope objectives. The result found by Dawes was, Least Angular Separation

[4] F. Twyman, *Prism and Lens Making* (Adam Hilger Ltd., London, 1942), pp. 119.

[5] A. E. Conrady, *Applied Optics and Optical Design* (Oxford University Press, New York, 1929), pp. 133.

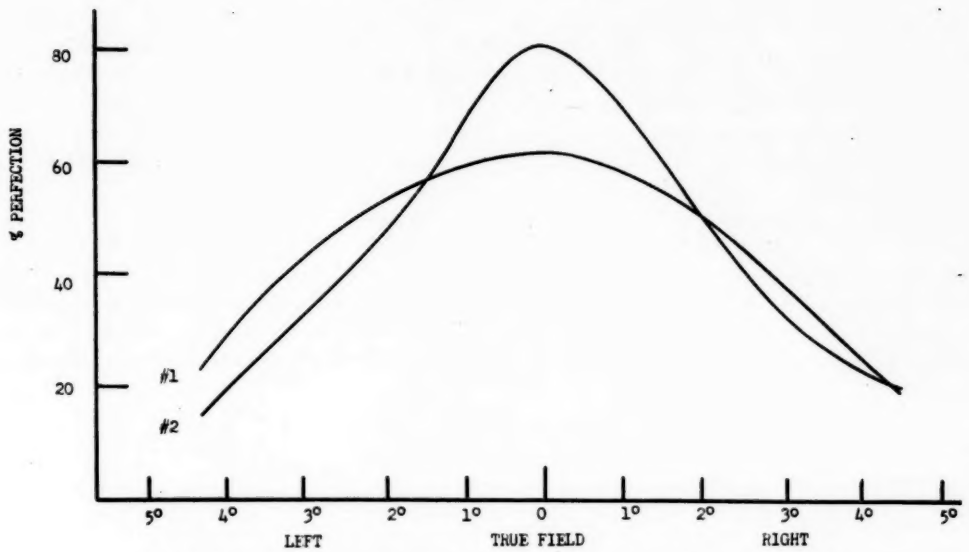


FIG. 2. Curve No. 1, average resolution characteristics for telescope with 3 different spherical erectors in conformance with design. Curve No. 2, average resolution characteristics for telescope with 3 different aspheric erectors. (Crown figured, 1 ring S.A.)

Resolvable = 4.5 seconds/aperture in inches. In spite of this work many books on optics still quote the classic formula, L.A.S.R. = 5.5 seconds/aperture in inches, a result surpassed by any telescope objective possessing a moderate degree of perfection.

Stars are practically never used as test objects for resolution measurements for obvious reasons. The author has measured the resolution of numerous objectives whose path error was known to be considerably less than $\lambda/4$ as determined with the Twyman interferometer. The results of these experiments yield L.A.S.R. = 4.75 seconds/aperture in inches for the Foucault test object. When the lengths of the bars and spaces on the Foucault test object are reduced to a value equal to half the grating space, the equation becomes L.A.S.R. $\sim$ 5.2 seconds/aperture in inches. Qualitatively both these results are to be expected when one considers the nature of the diffraction pattern produced by the Foucault target compared to the star pattern.

Figure 2 is a resumé of the resolution tests on a telescope using spherical erectors in conformance with design and erectors which had been figured to substantially remove the axial spherical aberration. Curve No. 1 is the average resolution *vs.* true field curve for a telescope with 3 different erecting lenses in conformance with design (S.A. app. 6 rings). Curve No. 2 is the average resolution *vs.* true field curve for the telescope with 3 different aspheric erectors having the crown element figures so that the residual spherical aberration was 1 ring. The ordinates (percent perfection) are the percentages of the resolution which was obtained compared to what could be obtained on the axis for a perfect objective of the same aperture. Curve No. 2 shows that the figured lenses produced resolutions in the telescope over 80 percent of perfection. This improvement was sufficient to reduce the axial "flicker" to an unobservable amount. Achievement of much more than 90 percent perfection in a lens-erecting telescope is rather uncommon even when the design is free of axial spherical aberration and has been properly executed. Mounting of the individual lenses in the telescope barrel cannot in general be accomplished so that each lens is exactly "squared on." As the result of these difficulties the resolution curves in general are slightly asymmetric. Curves No. 1 and 2, Fig. 2 show such asymmetry. The above-mentioned curves also tend to show that the resolution shown in

the off-axis imagery is slightly worse with the aspheric lenses than the spherical ones for large fields. These differences are small and approach the magnitude of the experimental error made in taking the off-axis measurements.

The figuring of moderately small lenses to remove primary spherical aberration amounting to approximately 1 percent of the focal length is not a difficult operation. It might be possible to design instruments with very desirable characteristics making use of the additional variable of an easily produced aspheric surface in conjunction with the normal spherical surfaces.

The author is indebted to Mr. Jos. Vrabel of the Frankford Arsenal Instrument Division for making the resolution measurements on these telescopes. The resolution measurements were made at The Pennsylvania State College under contract No. OEMsr-1197 O.S.R.D.-N.D.R.C. Section 16.1-optical instruments. The resolution data has been taken from Laboratory Report No. 11. The author wishes to express his appreciation for the many helpful discussions with Mr. O. K. Kaspereit and Mr. F. B. Patrick of Frankford Arsenal Design Division concerning some of the topics treated in this paper.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA VOLUME 36, NUMBER 5 MAY, 1946

Letter to the Editor

Trichromatic Units

PARRY MOON, *Massachusetts Institute of Technology, Cambridge, Massachusetts*
AND
DOMINA EBERLE SPENCER, *Tufts College, Medford, Massachusetts*
March 18, 1946

DR. JUDD's recent letter[1] to the Editor clarifies his position on trichromatic units. The letter, however, indicates a misconception that we feel should be rectified.

The statement is made that "the unit proposed by Moon and Spencer, to be known as the *young*, refers only to illuminance." On the contrary, our proposed units for X, Y, and Z refer to weighted flux densities at any point in space: they are not restricted to incident flux density or to radiated flux density. This view was explained[2] in 1942, and we have conformed to it ever since.

It is a peculiarity of lighting terminology that the idea of flux density is usually split into two separate concepts,[2] one indicating flux *to* a surface and the other flux *from* a surface. The C.I.E. adds a third concept, the generalized flux density, which is the one we have used. Since all three have the same dimensions and are expressed in what is really the same unit (lumen per unit area), though disguised under different names, it seems an unnecessary complication to use three concepts and three symbols where one suffices. This unnecessary complication has occurred in the lumen per unit area of photometrics. It has been imitated sometimes in the watt per unit area of radiometrics. But it should not be extended to colorimetrics, where it would give nine concepts where today there are only the three, X, Y, and Z. Yet such an extension seems to be implied by Dr. Judd when he suggests that we are restricting the *young* to incident flux density.

Since both *lux* and *footcandle* are units of incident flux density only, we are not at liberty to use either word for generalized flux density. Hence the introduction of the name *young* is not "piling confusion on confusion" but is a step toward a much-needed simplification.

The statement is also made that "Y may refer to luminous reflectance, luminous transmittance, illuminance, or luminance." That is true only when the defining equation is used in the loose way ordinarily employed by colorimetrists. If $J(\lambda)$ is in watt m^{-2} $micron^{-1}$, as stated in our letter,[3] then Y must be in weighted watts per unit area and cannot possibly be reflectance or transmittance, both of which are dimensionless. Also, since the concept $J(\lambda)$ was not restricted to incident or to radiated power, the Y of our equation is not restricted. The concept defined by the equation $Y=\int_0^\infty \bar{y}(\lambda)J(\lambda)d\lambda$ of our letter[3] is a definite quantity and should have a definite unit, for which we have suggested the name *young per square meter*.

An omission in our letter should be noted. In the sixth line from the end, *königs*, *youngs*, and *priests* should of course be königs/m^2, youngs/m^2, and priests/m^2.

We are grateful to Dr. Judd for the clear statement of his position, also for the emphasis that his letter inadvertently gives to the difference between the usual methods of the colorimetrist and the more precise formulation that is necessary for light sources.

[1] Deane B. Judd, J. Opt. Soc. Am. **36**, 120 (1946).
[2] Parry Moon, J. Opt. Soc. Am. **32**, 350 (1942).
[3] Parry Moon and Domina Eberle Spencer, J. Opt. Soc. Am. **36**, 120 (1946).